Carnegie Mellon University
Mellon College of Science

Thesis
Submitted in partial fulfillment of the requirements for the degree of Ph.D. in Biological Sciences

Title Improving 2D Gel Proteomics with the Structured Illumination Gel Imager

Presented by Phu T. Van

Accepted by the Department of Biological Sciences

Major Professor Jonathan S. Minden, Ph.D **Date** 5-5-2014

Department Head Nathan Urban, Ph.D **Date** 5-5-2014

Approved by the MCS College Council

Dean Fred Gilman, Ph. D **Date** 5-5-2014

Improving 2D Gel Proteomics with the Structured Illumination Gel Imager

Phu T. Van

Thesis Advisor: Dr. Jonathan S. Minden, PhD

for Mom and Dad

“Where there is much light, the shadows are deep” –Goethe

ACKNOWLEDGEMENTS

It is nearly impossible to thank everyone who made my doctoral work possible and my time in Pittsburgh enjoyable. But I will try.

Thank you, first and foremost, to my adviser, Jonathan Samuel Minden, whose scientific rigor, creativity, kindness and overwhelming tenacity will inspire me for years to come.

To Dr. Frederick Lanni, thank you for your humor, humanity and utter devotion to science and education. To Dr. Marcel Bruchez, thank you for your career advice and tips on public speaking. To Dr. James Schneider, thank you for agreeing to be on my committee, and for your kind words.

To Mom and Dad, thank you for the gift of life and your unwavering support. Dad's passing was my dark night of the soul, but your life pointed the way to the oncoming dawn.

To Vinitha Ganesan, Emily Furbee, Melissa Witzberger Krajcovic and Ashita Magal: thank you for your friendship in the Minden Lab, the long hours were made short by your company.

To the Minden undergraduates: Victor Bass, Amritha Parthasarathy, Alex Rodriguez, Danielle Schlesinger, Raghu Avula, Dagney Cooke, Minh Le, Alex Hurley, Travis Lear, Lawton Tellin, Rachel Willen, Taylor Maggiacomo, Anna Pyzel, thank you for your untempered passion for science, and for constantly reminding me how old I am.

To Adam Foote, Dan Shiwarski, Brendan Redler and Lina Song: thank you for bringing enthusiasm, creativity and insight to the Minden Lab, and for tolerating my often-intolerable jokes.

To Suchitra Ramachandran, Ritika Tewari, Stephanie Lewis *nee* Hughes and Michael Gamalinda: thank you for your company, advice and support through thick and thin.

To my classmates: thank you for making first year not only survivable, but pleasantly memorable, particularly John Pettersson and Jigar Desai: thank you for your friendship and making my endless walks around the second floor of Mellon Institute enjoyable.

To V. Emily Stark and Ena Miceli: thank you for your humor and helping me navigate the treacherous waters of regulations at department, university, city, state, global, and possibly galactic levels.

To Nitin Baliga, my former mentor at ISB, thank you for galvanizing my love for science, for giving me opportunities rarely afforded by mere interns, and faith on my long road to graduate school.

To Chris Loggers, my former mentor at USFS, thank you for your wisdom, for long conversations at unreasonable hours, for testing my abilities to explain science, and for reminding me that our highest achievement is kindness to one another.

To Scott and Tara LeDuc: thank you for being my home-away-from-home, your ever-present optimism and comfort were always the highlights of my trips back west.

To Achara, Wayne, Nancy, and Angela: thank you for your continued kindness and friendship

To Sara Thorson, Ivar Thorson and Filip Tkaczyk: thank you for your friendship, and showing me that family does not merely include those with common last names.

To Annie Lu: thank you for the trans-America adventure, exposure to questionable music, chicken, and for The Prince, a crazy idea brought to fruition through your talents and tireless devotion.

To Karlyn Beer and Ed Gano: thank you for our adventures in Truckie, interstate recipe exchange, cooking over Skype, and for reminding me that true friendships are not subject to distance.

To everyone who supported The Courtyard Kitchen: thank you for sharing your knowledge.

Thank you, everyone, and I hope each day will bring you another measure of happiness, whatever you imagine it to be.

ABSTRACT

This thesis is composed of three separate projects:

1. Proteomics is the study of complex protein mixtures found in a cell, organ, or entire organism. The vast concentration range of these samples, estimated at approximately 150,000-fold for simple unicellular eukaryotes is beyond current detection methods. We present a technology called Structured Illumination (SI) Gel Imager that employs an LCD projector to selectively illuminate fluorescently-labeled proteins separated into individual protein spots on 2-dimensional electrophoresis (2DE) gels. SI Gel Imager images have a dynamic range of approximately 1,000,000-fold, making it a valuable tool for proteomic detection.
2. 2DE gels possess the ability to separate proteins with extremely high resolution of molecular-weight and isoelectric-point. However, they suffer from variable sample loss incurred during protein reduction and alkylation steps required for subsequent sequencing by mass-spectrometry, up to about 30% of the starting sample. We present a protein equilibration method utilizing agarose stacking gels to reduce experiment variability and sample loss.
3. 2DE-based proteomics is a time-consuming process, requiring up to 3 days, and suffers from low reproducibility. To provide undergraduates to the experience of conducting proteomics research, we developed the Proteomics Platoon approach, where a group of undergraduate students work in two-person teams to perform proteomic experiments using a wide variety of biological samples. The close-knit nature of the Platoon further fosters collaboration, communication and mentorship while completing complex

scientific projects. The Platoon approach serves as a model for involving undergraduates in complex research projects.

TABLE OF CONTENTS

LIST OF FIGURES

Introduction

Chapter 1

LIST OF TABLES

INTRODUCTION

Proteomics - the study of the entire protein complement of a tissue, cell, or organism - has generated enormous interest as a logical progression of genomics. As the ultimate effectors in the cell, changes either in protein abundance or physical characteristics often have dramatic physiological consequences. A major aim for proteomics therefore has been to create “atlases” for the proteome of humans and other model organisms, from which we can glean new physiological insights and discern markers of disease.

However, proteomics has been plagued by the complexity of the protein mixtures being studied, consisting of many protein species over a vast concentration range. For example, in the budding yeast *Saccharomyces cerevisiae,* proteins have been measured to range from about ten copies per cell to ~1.5 million copies per cell, giving an approximately 150,000-fold concentration range (Ghaemmaghami et al., 2003). In humans, serum protein concentrations have even broader ranges. The extraordinary biochemical diversity of proteins has so far resisted our efforts to easily amplify (or specifically enrich for) scarce proteins the way polymerase-chain-reaction could amplify nucleotides. As a result, these large concentration ranges pose an important hurdle for unbiased, proteome-wide comparisons that search for changes in protein concentration and post-translational modification of all cellular proteins.

Current comparative proteomics strategies fall into two general categories referred to as “peptide-centric” and “protein-centric”. The peptide-centric approach relies on mass spectrometry (MS) for protein quantification by the detection of peptides arising from enzymatic digestion of protein samples (Aebersold and Mann, 2003). This commonly used method employs liquid chromatography followed by tandem MS (LC-MS/MS) to detect peptides over a ~1,000-fold concentration range (Stahl-Zeng et al., 2007). Selected Reaction Monitoring (SRM) is an

MS refinement that utilizes a list of user-defined proteins to direct the search and quantification of selected peptides and their fragmentation products. While SRM is a powerful quantification tool, it requires *a priori* knowledge of which specific fragmentation reactions may occur (Lange et al., 2008). Therefore, SRM cannot be used for unbiased, comparative proteome studies. While MS-based proteomics has proven to be very useful and has consequently been widely adopted in many experimental settings, tryptic digestion of the protein samples remains a fundamental difficult complication. To illustrate this we obtained the sequence for 1,000 random proteins in *Saccharomyces cerevisiae*, and then performed an *in silico* digest using the program Proteogest (Cagney et al., 2003) with the default parameters generated 28,212 distinct peptides.

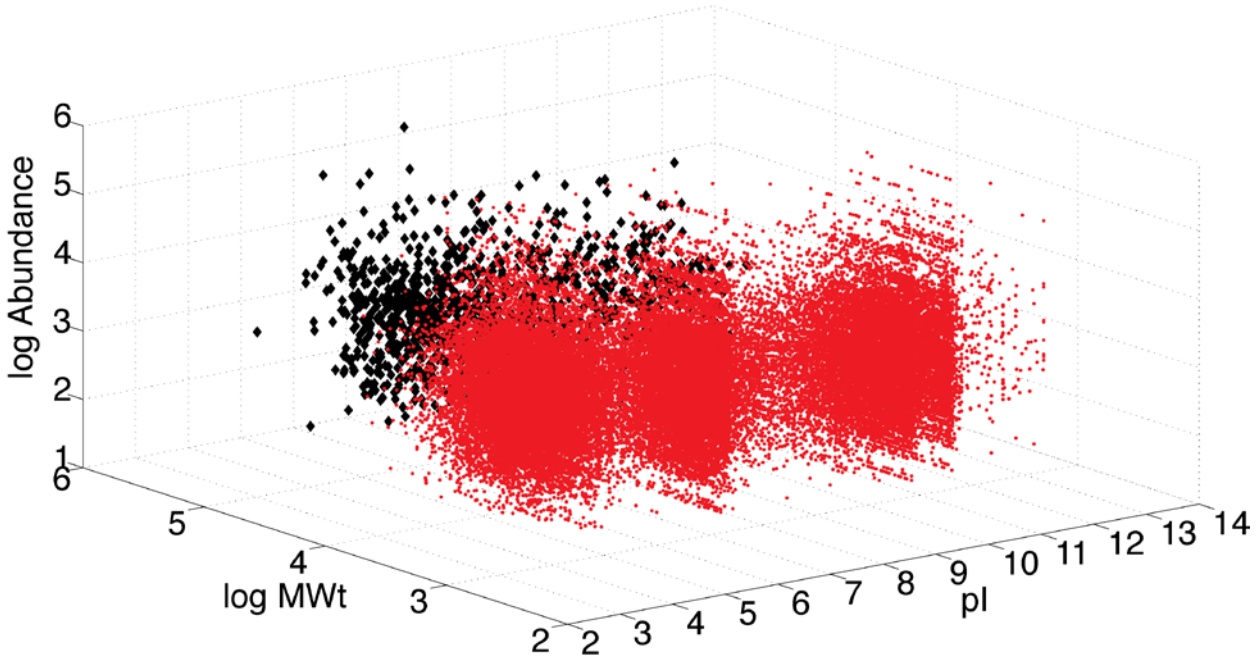


Figure 0-1. *The chemical space of 1,000 random yeast proteins and their tryptic peptides. Black dots show randomly-selected proteins by their molecular weight (MWt), isoelectric point (pI) and abundance according to Ghammagaemmi et al. Red dots show the tryptic peptides of these proteins as generated by the program Proteogest. The marker size for the red dots has been reduced as to not crowd the image.*

As demonstrated in Figure 0-1, these tryptic peptides greatly increased the complexity of the mixture being analyzed, necessitating additional processing steps such as fractionating the peptide mixture and analyzing each fraction separately. The large number of tryptic fragments also complicates identification, since each fragment needs to be unambiguously mapped back to its respective parent protein. Though MS is exquisitely sensitive, the increasing number of tryptic fragments makes unambiguous identification of protein changes in complex mixtures very difficult and time-consuming. For these reasons, neither traditional LCMS nor SRM in their current incarnations are suitable for unbiased high dynamic range detection of complex proteomic samples.

The protein-centric approach divides this comparative proteomics method into two main tasks: (1) separation of intact proteins by 2D gel electrophoresis (2DE) followed by quantitative protein detection and (2) identifying proteins of interest by MS, as shown in Figure 0-2.

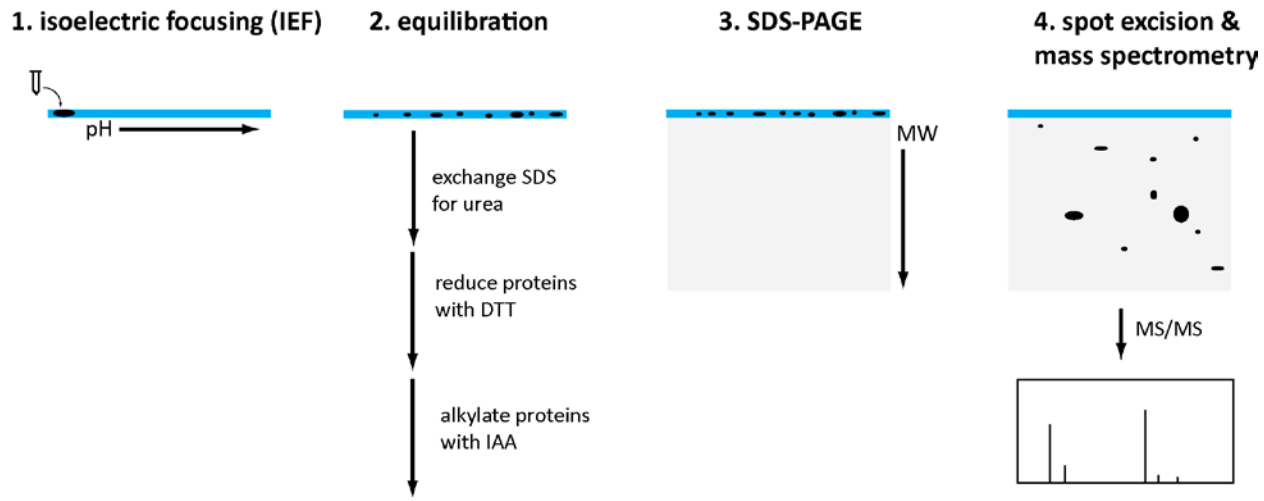


Figure 0-2. *The typical workflow of 2DE-MS/MS experiments. Protein samples are first separated by isoelectric point using IEF, equilibrated, then separated by molecular weight using SDS-PAGE. The final step involves protein excision from SDS-PAGE gel and sequencing using LCMS/MS*

Two-dimensional electrophoresis (2DE) in its current form was first reported independently in 1975 by O'Farrell (O'Farrell, 1975) and Klose (Klose, 1975). These efforts resulted in improvements over the older method of 2D maps which used native isoelectric focusing (IEF) and native PAGE ((Macko and Stegemann, 1969), (Dale and Latner, 1969) and others). O'Farrell's protocol introduced detergents and chaotropes for better protein separation, resulting in a dramatic improvement in gel quality. Since then, 2DE has attracted the attention of many researchers and gained steady improvements in spatial separation and sensitivity. In terms of spatial separation, gradient gels with variable porosity came in 1984, which allowed for better separation in the second dimension (Walker, 1984). The invention of immobilized pH gradients (IPG) came in 1993 which improved IEF for first dimension separation of proteins (Bjellqvist et al., 1993), overcoming limitations of pH gradients generated using only carrier ampholytes. The later development of narrow pH-range IPG strips has greatly increased the resolution of 2DE gels while reducing overlapping spots. Current 2DE experimenters benefit from extensive commercial development of these early discoveries: gradient gel mixers, computerized IEF devices, precast large-format gels, IPG strips and associated buffers and reagents.

At the same time, considerable progress has also been made in labeling the protein sample for visual detection. The comparatively simpler approach is to stain gels after electrophoresis, where the dyes permeate through the gel matrix and attach to the immobilized proteins. Particularly well-known is Coomassie Brilliant Blue (CBB), a dye used in rudimentary forms since the 19th century and first applied to stain proteins in 1963 (de St Groth et al., 1963). Silver staining was publicized in 1973, which afforded much greater sensitivity compared to CBB (Kerenyi and Gallyas, 1973).

Proteins can also be labeled directly by attaching genetically-encoded fluorophores (green fluorescent protein (Shimomura et al., 1962) or variants) or chemically attaching small-molecular-weight dyes, eliminating the need for staining the entire gel. Difference Gel Electrophoresis (DIGE) is a commonly used method for comparing different fluorescently-tagged proteome preparations (Unlu et al., 1997). DIGE was made possible with the synthesis of mass- and charge-matched fluorophores, themselves based on a well-researched family of cyanine-derived dyes (Ernst et al., 1989).

Standard 2DE gels are capable of resolving ~2000 protein spots per gel; large-format 2DE gels can resolve more than 10,000 spots (Zabel and Klose, 2009). 2DE gels have very high resolution, capable of resolving spots different by 0.001 pH units; and are also extremely sensitive, able to detect down to <1 ng of protein per spot (Görg et al., 2004). This high sensitivity and spatial separation make 2DE gels well-suited not only for detecting changes in protein abundance, but also for distinguishing between different protein isoforms and post-translational modifications. Limitations of gel-based proteomics include co-migration of proteins leading to overlapping gel spots, difficulty with resolving very large or very hydrophobic (e.g. membrane-integral) proteins(Adessi et al., 1997), and low reproducibility due to inter-gel variation (Voss and Haberl, 2000). Some advances like lipid removal, deglycosylation and the creation of novel detergents have made detecting membrane proteins more manageable (Molloy et al., 1998; Speers and Wu, 2007). Therefore, we believe that 2DE gels are a promising starting point to improve proteomic experiments.

As described in this thesis, my doctoral work set out to improve gel-based, protein-centric proteomics in three ways. First, we sought to improve fluorescence gel imaging. Quantification of protein spots in 2DE gels is done either with scientific-grade cameras using

full-field illumination, or laser scanning imagers. In the resultant images, highly abundant proteins appear as large, bright spots; less abundant proteins appear as small, dim spots. The development of DIGE allowed the running and thus comparison of two or three samples in the same gel using mass- and charge-matched cyanine-based, fluorescent dyes (Unlu et al., 1997). Since the samples are subjected to the same electrophoretic forces simultaneously, inter-gel variability is eliminated. DIGE allows for the detection of very small differences in protein abundance, charge or mass (as low as ± 15%). DIGE has been used in many studies to obtain comprehensive protein "snapshots" of complex samples (Minden et al., 2009). Previous studies estimated that full-field and laser scanning imaging of 2D DIGE gels has a dynamic range of about 10,000-fold, defined as the ratio between highest and lowest fluorescent intensity of detected spots (Minden, 2007). While promising, this performance still falls short of the dynamic range required for whole-cell proteome studies. The dynamic range of 2DE gel images is limited by the camera used, the most common models having a 16-bit detector.

In response, we applied structured illumination to extend the dynamic range of 2DE gel images; implemented in a device we called the Structured Illumination (SI) Gel Imager. SI Gel Imager selectively illuminates parts of a fluorescent gel while acquiring luminance data by integrating increasingly long exposures to increase effective image dynamic range.

The second improvement of my thesis focuses on 2DE gels' sample retention. Most 2DE experiments involve a final protein identification step in which the proteins of interest are excised for the gel and sequenced using mass spectrometry (MS), either Matrix-Assisted-Laser-Desorption-Ionization-Time-of-Flight (MALDI-TOF) or high-performance liquid-chromatography tandem MS (LCMS/MS) (Aebersold and Mann, 2003). Many proteins contains cysteine residues, many linked in pairs to form disulfide "bridges", a major protein structural

feature. During 2DE, these disulfide bridges can break, and reform unpredictably to connect polypeptide chains that were previously unconnected, complicating protein sequencing. As a result, sequencing via MS requires sample equilibration step between the first and second dimension of 2DE (Herbert et al., 2001), where a reducing agent is applied to the protein sample to prevent the formation of disulfide bridges. Originally β-mercaptoethanol (βME, BME, or 2ME) was used, but it has since been displaced by dithiothreitol (DTT), a more powerful reducing agent (Cleland, 1964). The second part of equilibration is the permanent "capping" of the free –SH groups by an alkylating agent like iodoacetamide (IAA). The equilibration process also exchanges the protein denaturants used in first dimension IEF, most commonly urea or CHAPS with SDS used in the second dimension SDS-PAGE. Taken together, reduction and alkylation has been shown to improve the quality of 2DE gels, both with respect to the elimination of spurious cross-linked polypeptide chains, and reducing "streaks" in the resultant 2DE gels (Gorg et al., 1987). While it has been adapted as part of the standard 2DE workflow, current sample equilibration suffers from variable sample loss. The IPG strips are stirred or shaken in equilibration solutions containing the appropriate concentrations of DTT and IAA, allowing the protein samples immobilized in IPGs to diffuse out and thus become lost to latter parts of the 2DE workflow. In response we propose an equilibration procedure in which the protein sample is electrophoresed through agarose stacking gels containing SDS, DTT and IAA immediately before 2nd dimension SDS-PAGE. We show that this approach achieves SDS exchange, reduction and alkylation of the existing equilibration method while reducing sample loss and improves gel-to-gel variability.

The third element of my thesis focuses on furthering undergraduate education, particularly peer mentorship and collaboration. While it is common for many laboratories

including ours to employ undergraduate students on a part-time basis, proteomics experiments using 2DE gels are time-consuming, due to the long time required for separation by IEF and large format SDS-PAGE. While many reagents have been made standard and commercially-available and some attempts have been made at automating 2DE (Hiratsuka et al., 2007), for the time being, running 2DE gels remain a time-consuming, complex, multi-step procedure incompatible with the hectic schedule of a typical undergraduate student. Another concern in employing undergraduate students is the students' intellectual development. In most laboratories, undergraduate students are often assigned to a single project for the duration of his or her internship, guided by a post-doctoral fellow or graduate student with little chance for sustained collaborative work in small groups.

To address these problems, we instituted a refinement of undergraduate internship in our laboratory that we call "the platoon system". The undergraduate students in our lab are organized into teams of two, a senior student paired with a junior student. After several semesters of working together, the senior student takes on the responsibility of mentoring the junior student, under the overall guidance of the faculty member and graduate students.

This system encourages leadership in the senior student, provides accessible peer instruction to the junior student and promotes communication and collaboration in both partners. As a whole, the platoon benefits from the common shared experience of running (and therefore troubleshooting) 2DE gels and provide mutual time coverage for experiments using diverse biological samples obtained from collaborators. After two and a half years of implementation, we have gained improvements in 2DE experiments, fostered a realistic undergraduate research experiences while helping our undergraduate students develop useful scientific and collaborative skills.

CHAPTER 1: THE STRUCTURED ILLUMINATION GEL IMAGER (S.I.G.I.)

Most of the work described in this chapter was submitted for publication. A major limitation to detecting the full concentration range of cellular proteins in 2DE gels is the dynamic range of the imaging device. In terms of fluorescence imagers that use full-field illumination and Charged-Coupled Device (CCD) cameras, abundant protein spots are highly fluorescent and thus require relatively short exposures. Low abundance protein spots require longer exposures. 2DE gels often have high abundance proteins in close proximity to low abundance proteins. Exposure times that are required to detect low abundance proteins typically lead to pixel saturation due to the fluorescent signals arising from high abundance proteins. Saturated pixels all report the detector's maximum detection value, regardless of their true intensity. The detectors employed for fluorescence imaging usually utilize 16-bit CCD cameras or photomultipliers, thus saturating at 65,535 (2^{16}-1) counts. Given the noise characteristics of electronic detectors, as well as the variability of biological samples and significant sources of background fluorescence signals, the current detection-range limit of fluorescent gel imaging systems is about 10,000-fold (Marouga et al., 2005). Since quantification of protein spots depends on accurate intensity values for all pixels within a gel image, overcoming this detector saturation limit is essential for obtaining high dynamic range gel images.

Here, we describe the development of a Structured Illumination (SI) gel imager, a fluorescence imaging technology that overcomes the pixel saturation limit of current fluorescence gel imagers. SI is a broad term meaning that different parts of the imaging area are illuminated with different amounts of light. SI has been extensively studied using different techniques towards different goals, with applications in robot guidance (Srinivasan and Lumia, 1988), remote sensing (Cortizo et al., 2003; Erkmen, 2011), 3-dimensional surface scanning

(Cortizo et al., 2003) and document processing (Bertucci et al., 2003). There are also notable SI applications in microscopy in which SI is employed for axial and transverse superresolution (Zeiss Apotome2, Leica OptiGrid). To date, no SI application involves the imaging of electrophoretic gels. In the application of structured illumination described here, a video projector is used to create a series of binary masks that progressively block the illumination of high abundance proteins as the exposure time is progressively increased to detect low abundance proteins. This scheme produces a series of images with increasing exposure times, and avoids pixel saturation. The SI gel images are assembled into a single high-dynamic-range image. Since saturation never occurs, all measured (un-masked) pixel intensity values remain reliable, allowing for the detection of proteins over a 1,000,000-fold concentration range in a single field of view (FOV) of an SDS-PAGE gel and significant increase in the number of protein spots detected in a 2DE gel containing whole extracts.

MATERIALS AND METHODS

The Structured Illumination Gel Imager

The central principle of the SI gel imager is to create a feedback loop between the imaging camera and a video projector acting as a light source. To initiate this loop, an initial short exposure (*N* seconds: *N* is set so that no image pixels are saturated) of a gel is recorded using uniform, full-field illumination from the video projector. Any camera pixels registering a signal above a pre-determined threshold, which in this case is 30,000 counts (just under half of the pixel saturation limit for a 16-bit camera), will be masked in subsequent exposures. A ‘masked image’ of the gel is computed by turning off (setting to black) all projector pixels above the threshold; the rest of the projector pixels are set to full on (white). This masked image is

projected onto the gel and a 2*N* exposure is captured. The new image is used to compute the next masked image that is projected onto the gel and a 4*N* exposure is recorded. Once a pixel is masked, it remains masked throughout the imaging series. Typically, this cycle is repeated 6-8 times, with a maximum exposure time of 64 seconds. These parameters ensure that none of the pixels in the series of SI images ever become saturated. By assembling intensity data from this sequence of masked images, the SI gel imager achieves an improved dynamic range over conventional full-field illumination, single-exposure imagers.

The masking strategy is simulated in Fig. 1 using a typical 2DE gel image. In Fig. 1-1A, the unmasked 16-bit image has readily visible protein spots. The brightest, most abundant protein in this field is indicated by the red arrow and one of the dimmest proteins is indicated by the green arrow. A binary mask can be generated by setting all pixels exceeding a preset intensity threshold (30,000 in this example) to zero, represented as black, and unmasked pixels are set to one, shown as white (Fig. 1-1B). When this mask is applied onto the original gel image by image multiplication, bright spots are masked but many previously unseen, dim spots become visible relative to the dimmest protein spot from the unmasked image (Fig. 1-1C). In Image A, obtaining a non-saturated image of the brightest spot (red arrow) with uniform illumination limits the dim spot (green arrow) to low dynamic range. In a subsequent masked image, the dimmest spots, like the spot indicated by the green arrow, can be boosted to high dynamic range.

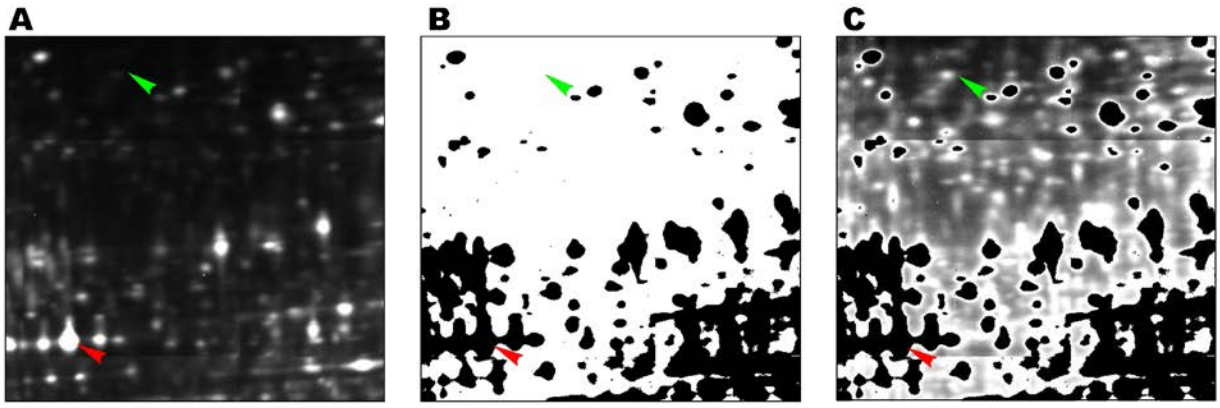


Figure 1-1. *Structured illumination improves quality of fluorescent gel images.* ***(A)*** *Fluorescence image of a typical 2DE protein gel. Abundant proteins appear as bright spots (red arrow) while low-abundance proteins are not visible (green arrow).* ***(B)*** *A binary mask generated from A: pixels above a certain threshold (30,000 in this example) were set to zero, shown as black; the remaining pixels were set to one, shown as white.* ***(C)*** *Gel image with mask overlaid. Shown here is the product of multiplying image A by the binary mask shown in B. The bright spot (red arrow) is now masked but many dim spots are now visible when the image is rescaled, including one previously unseen in A (green arrow). Notice that even dimmer spots are more evident after masking.*

Construction of the SI Gel Imager

The main components of the SI gel imager are illustrated in Fig. 1-2A. An auditorium-grade LCD projector, with its standard lens removed, was used as the SI light source (NEC Display of America, Itasca, IL). Light emitted from the horizontally-mounted video projector was focused through a pair of achromatic lenses to form an image of the video projector's LCD array onto the gel (f=200 mm, Edmund Optics, Barrington, NJ), then passed through fluorescence excitation bandpass filters mounted in a motorized filter wheel (CVI, Albuquerque, NM) and redirected onto the gel via a multi-bandpass dichroic mirror mounted at 45°. The gel was placed in a black-anodized aluminum tray with a bottom plate made from a very low-fluorescence, fused-silica window. To reduce back-reflected light, the fused-silica window was broad-spectrum, anti-reflection coated (BBAR) (ESCO, Oak Ridge, NJ). To further reduce stray light, excitation light that passed through the gel was collected by an L-shaped light trap made of poly(methyl methacrylate) that was painted flat-black on the interior. Fluorescence emission

from the fluorescently-labeled proteins in the gel pass back up through the dichroic beam splitter and matching emission band-pass filters mounted in a second, motorized filter wheel into the camera. The entire light path was enclosed in 2”-diameter flat black optical tubing to limit stray light. Most optical mounting components were from ThorLabs (Newton, NJ). The 2”-diameter optical filters for Cy3 and Cy5 were from Chroma Technology (Bellows Falls, VT). Optical elements were mounted on a three-axis stage and positioned with manual micrometers (Newport Corporation, Irvine, CA). The light-tight, SI gel imager housing was constructed from black anodized structural aluminum and black poly(methyl methacrylate) with the interior painted flat-black (80/20 Inc, Ft. Wayne, IN).

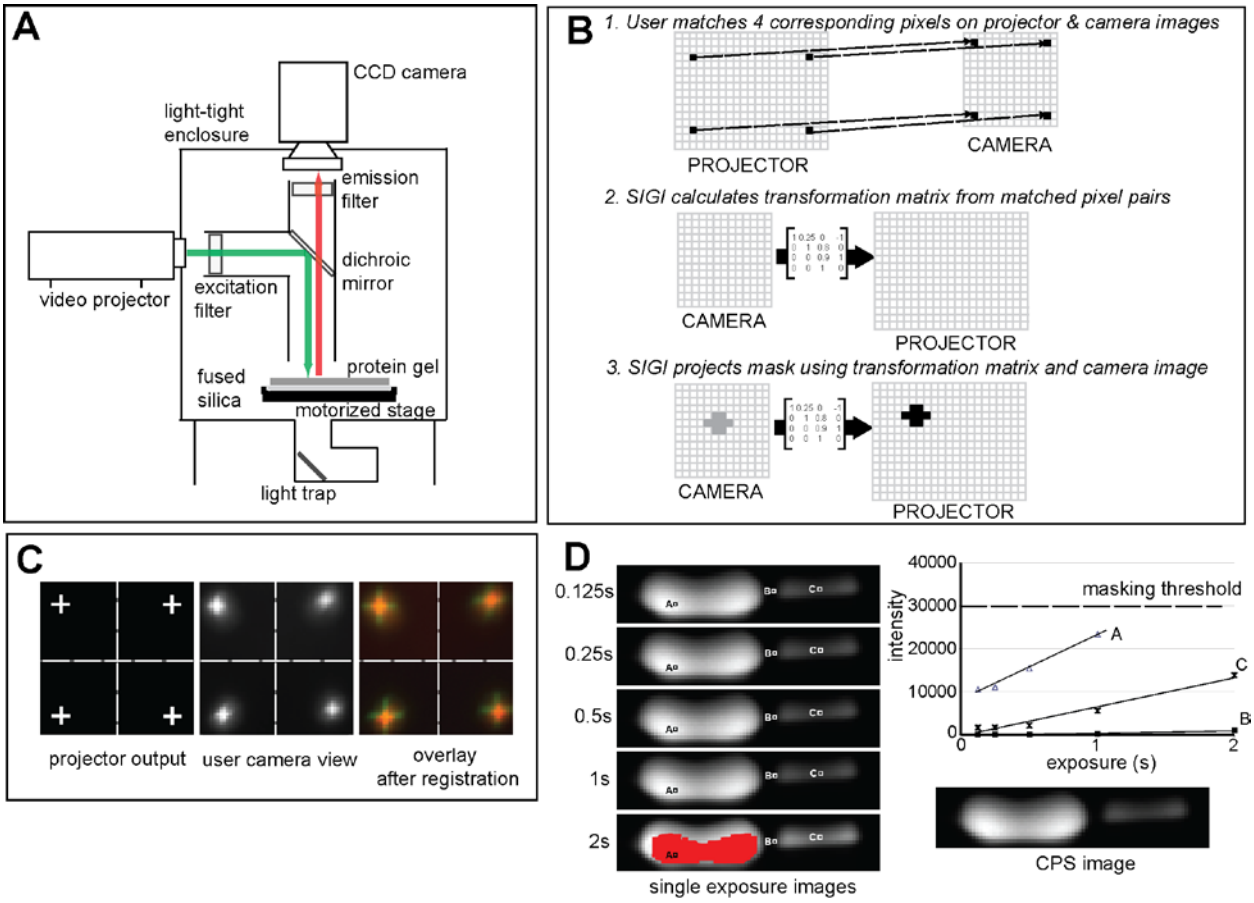


Figure 1-2. *Components and operation of the SI gel imager.* ***(A)*** *Diagram of the SI gel imager’s*

optical configuration. ***(B)*** *Workflow of the SI gel imager's image registration process.* ***(C)*** *Shown here is the image registration results where sub-images of the four corners of imager's FOV are displayed. After perspective transformation has been calculated, registration was confirmed by overlaying the projector output image (left panel) on camera image (middle panel). The resultant is shown as a false-colored image of overlapping crosses (right panel: red is the camera image, green is the projector output).* ***(D)*** *High dynamic range imaging with the SI gel imager: successively longer exposures were acquired, with masks (false-colored red) being projected when pixel intensity reached the masking threshold. For each pixel, a Counts-Per-Second (CPS) value was derived as the slope of the digital counts versus exposure time graph of unmasked exposures and stored as a 32-bit floating-point number in CPS images.*

The projector used had a native resolution of 1024x768 pixels and a manufacturer-rated ANSI contrast ratio of 1:1000. The camera was an actively-cooled 16-bit scientific CCD camera (Roper Scientific, Sarasota, FL) with a native resolution of 1300 x 1340 pixels. For the SI gel imager, the CCD's acquisition field was cropped to the central 1024x1024 pixels of the CCD and binned (4x4) to produce a 256x256 pixel image. These CCD adjustments ensure projector oversampling and increase S/N ratio of the camera image. To minimize projector spatial non-uniformity, the SI gel imager was aligned so that the center of the projection area was used for creating masks. Projector brightness and contrast settings were set at intermediate values to minimize "clipping", a phenomenon where very bright and very dim projector pixels are not displayed due to excessive contrast (Kwak and MacDonald, 2000). The gel plate accommodated 2DE gels up to 250 x 200 mm. At 0.164 mm/pixel, the camera FOV was 42x42mm which necessitated image tiling to completely capture large-format 2DE gel images. The gel tray was

loaded onto an XY, motorized stage that provided two-axis movement relative to the camera's fixed field of view with 1μm resolution (NEAT, Beverly, MA).

The SI gel imager is controlled by a dedicated computer workstation (Dell, Inc., Round Rock, TX) via a purpose-built graphical software package "SILab" that directs the imager's hardware components through serial port connections. SILab executes the imager's main operations: acquire images, generate image masks for the projector, direct stage movements for tiled acquisition, and perform rudimentary image operations such as zooming, intensity scaling and animation of two-frame DIGE movies. SILab also performs image registration and assembles the final high dynamic range images automatically after capture.

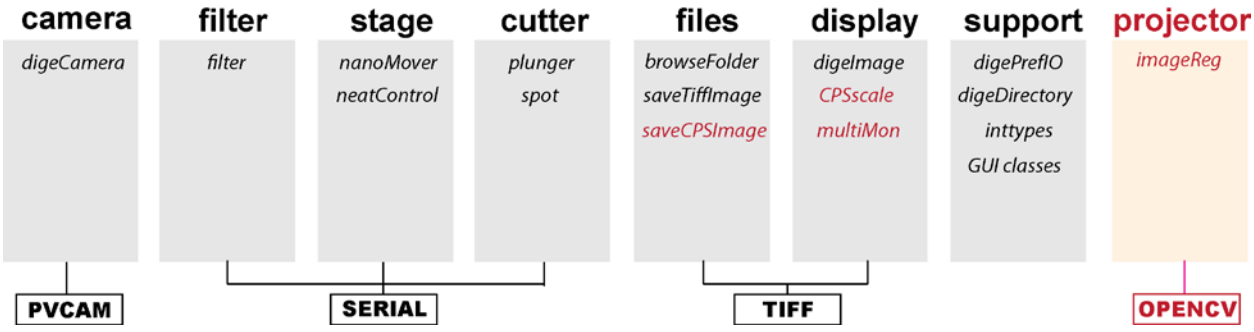


Figure 1-3. *Components of SILab software program as of March 2013. Functions ported from WinDige are in black, new functions are in red. Capitalized texts in boxes are external software libraries.*

SILab is an interactive C/C++ program written for Microsoft Windows with a graphical user interface, whose major components are shown in Figure 1-3. The program is derived from WinDige, a program originally written at CMU circa 1998 to control a similar imager under Windows 95. WinDige, and thus SILab makes use of several external software libraries: PVCAM for acquisition of raw image intensities; SERIAL for communication with filter wheels

and linear stages making up the gel stage and robotic gel cutter; TIFF for storing and displaying single-shot images (high-dynamic-range CPS images generated by SI Gel Imager is stored as uncompressed 32-bit floating point arrays, as discussed later under heading “SI Gel Imager Acquisition of High Dynamic Range Images”). Upon my involvement of the project, the WinDige codebase was rewritten for Windows 7 using then-current programming conventions and software development tools. The structured illumination functionality also required additional C/C++ code and external software libraries, most notably OpenCV, which enabled image registration (Figure 1-3, red text). During development, approximately 7,500 lines of code were changed, added or removed out of ~21,000 lines of code that made up SILab. The main user interface of SILab is shown in Figure 1-4.

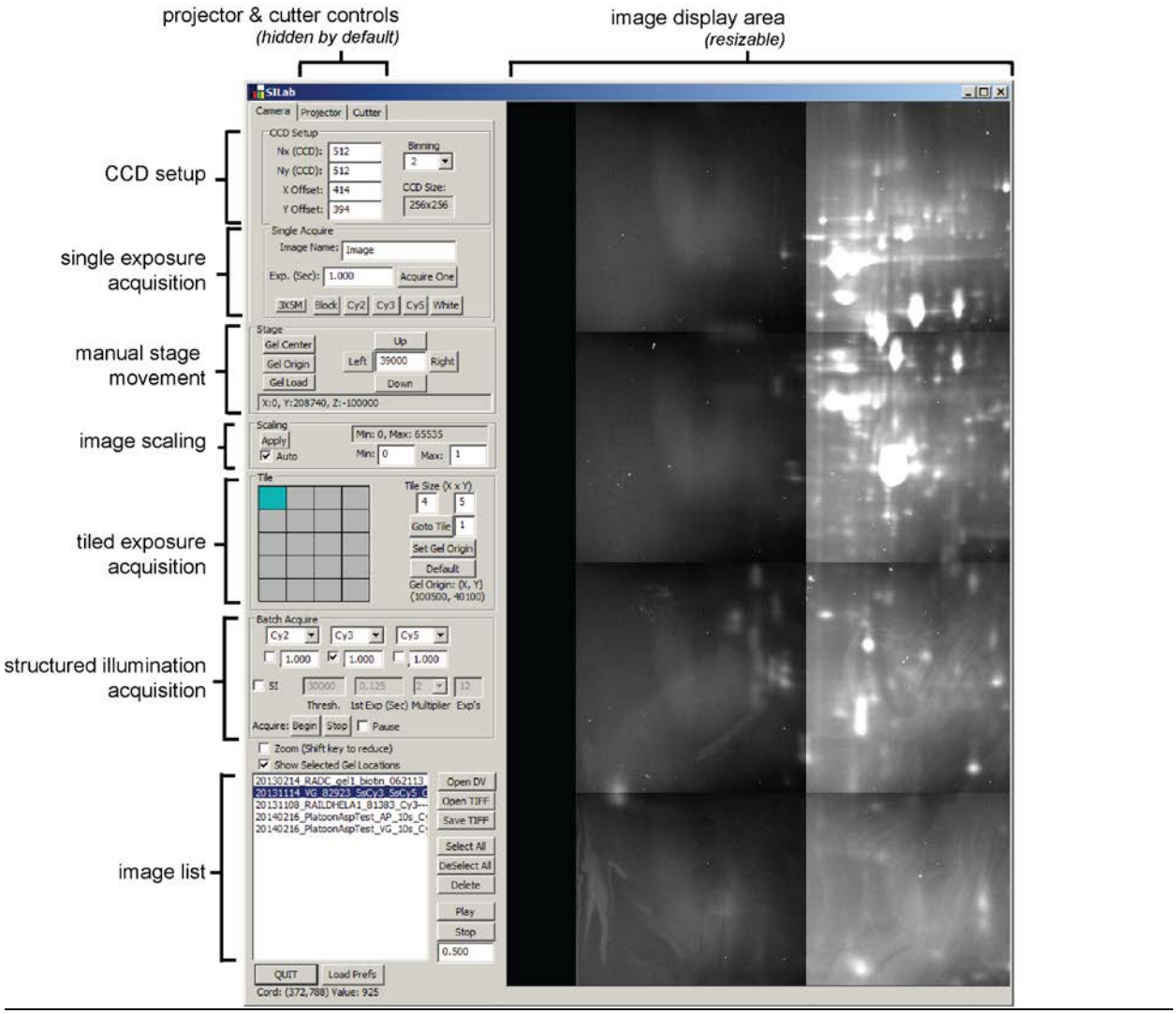


Figure 1-4. *Graphical user interface of SILab, as of March 2014.*

Camera-Projector Image Registration

Accurately projecting masks onto gel spots requires mapping every pixel of the projector image to its counterpart pixel in the camera FOV, a process called “image registration”. Many image registration methods exist to satisfy different constraints on registration speed, accuracy, resistance to image noise and automation (Zitova and Flusser, 2003). We require pixel-level alignment for the entire camera FOV with a relatively simple optical configuration (rigid planar

camera and projector surfaces, monochromatic camera and projector images) and imaging requirements (static camera and projector positions, low acquisition speed, oversampled projector image, where the projector a larger pixel array than the camera's pixel array). For the SI gel imager, a control point matching method was chosen for supervised image registration, outlined in Fig. 1-2B. Briefly, the user selects corresponding pixels in camera and projector images, after which these coordinates are used to calculate a transformation matrix that is used to map camera pixels into the projector image plane.

In SILab, this is implemented as follows: first, the projector casts images of crosses (5x5 pixel "+" signs) that are positioned at the four corners of the camera FOV onto a uniformly fluorescent target – a 1.5 mm thick, agarose slab gel containing Cy3-labeled protein (serum albumin) (Fig. 1-2C, left panel). An enlarged version of the resultant camera image is displayed on SILab viewer (Fig. 1-2C, middle panel). The user then clicks on the center of each cross with the mouse cursor, which registers the corresponding coordinates for the projector image. A projective transformation matrix between the two sets of coordinates is calculated:

$$[CCD_x, CCD_y] = M * [Proj_x, Proj_y]$$

, where $Proj_x$ and $Proj_y$ are x- and y-coordinate of the projector pixels, CCD_x and CCD_y are x- and y- coordinates of the camera pixels and M is the matrix for a perspective transformation. SILab calculates M using the OpenCV software toolkit (Bradski, 2000). OpenCV's getPerspectiveTransform() function takes $Proj_x$, $Proj_y$ from the projector image crosshair coordinates, CCD_x and CCD_y from the user clicks and generates M explicitly. The camera image is then transformed using OpenCV's warpPerspective() function using M for the transformation parameters. The combined image of camera crosses (shown in red) overlaid on projector crosses (shown in green) is presented to the user for confirmation of image registration (Fig. 1-2C, right

panel). Afterwards, the inverse matrix M^{-1} is automatically used by SILab to project masks for all subsequent images. This supervised process makes image registration a highly robust process.

Image registration is required infrequently, either when the projector or camera have been displaced for system maintenance or by minor vibrations during normal use. In our laboratory, SILab requires image registration approximately every 15-20 gels, about once a month of typical use, with each registration session taking 1-2 minutes.

Since image registration is fundamental to SI Gel Imager operation, we investigated other image registration methods before settling on control point matching. Initially for simplicity we tested light-striping and feature matching, both well-established image registration methods (ref). Light striping works by projecting an alternating black-and-white horizontal or vertical "zebra-striped" patterns of different thicknesses, creating a projector image stack (Figure 1-5, top row). We then project and record each image with the camera, creating a camera image stack (Figure 1-5, middle row). Each pixel in the camera and projector camera fields of view then has an identifying "barcode" which can then be computationally matched to each other, generating a lookup table (LUT, Figure 1-5, bottom row). For projecting masks onto the gel, each projector pixel is found through the LUT and turned on or off. This process is illustrated in Figure 1-5. Light striping is unsupervised, adequately fast and quite successful given sufficient image contrast. However, we find that image contrast significant degraded at the edges of the projector FOV, resulting in ambiguous barcodes and unregistered pixels. Since the SI Gel Imager acquires tiled images, unregistered edge pixels were unacceptable and we discarded light-striping.

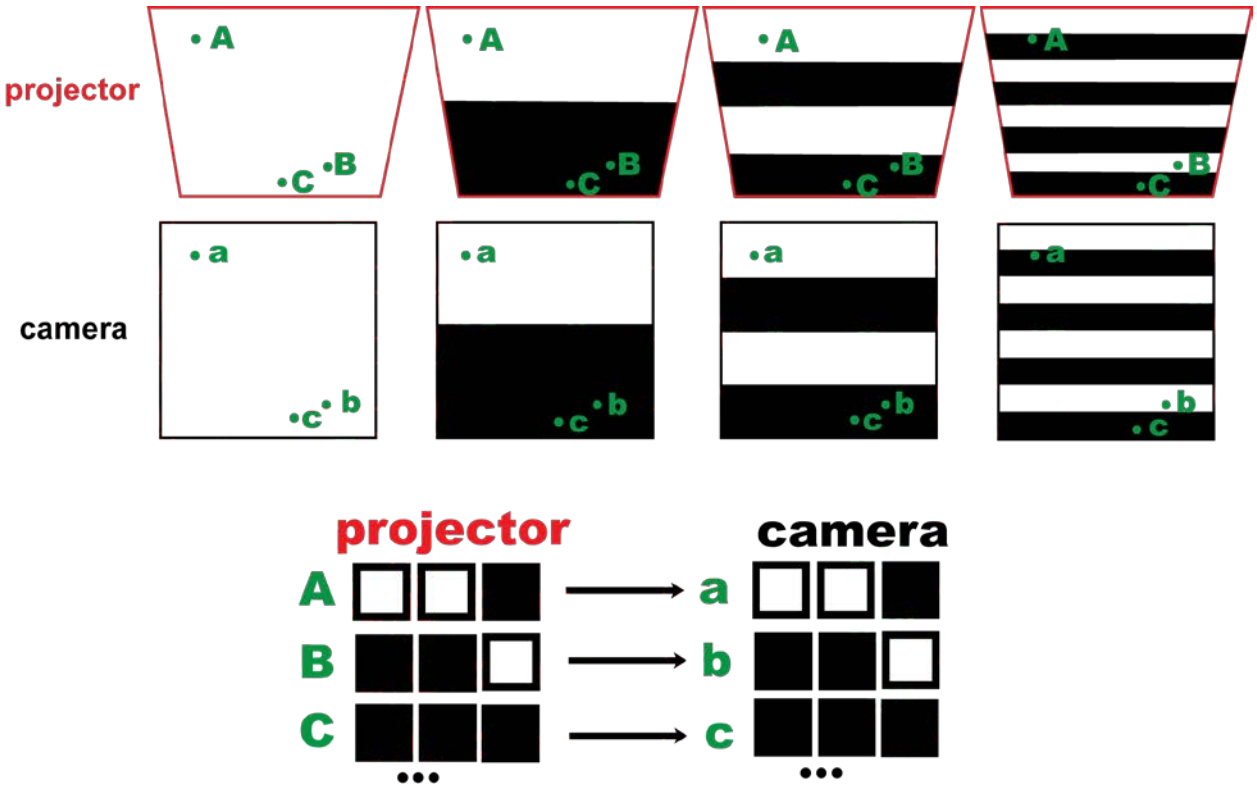


Figure 1-5. *Image registration by light striping. In this example, we are registering the projector image (red trapezoid) with the camera image (black square), matching projector pixels (A, B, C) to camera pixels (a, b, c respectively). Progressively thinner alternating "zebra-stripe" patterns are projected (top row), and imaged (middle row). The resultant projector and camera image stacks contain one projector and one camera "barcode" for each pixel, which can be computationally matched, achieving one-to-one correspondence between the two images (bottom row). For simplicity, the corresponding vertical "zebra-stripe" patterns have been omitted.*

Feature matching is another unsupervised image registration method we explored for the SI imager. A single grid of geometric shapes is projected and imaged. OpenCV then attempts to computationally match the projector and camera pixels automatically, iteratively searching for the necessary transformation matrix until converging on an optimized solution (Fig. 1-6).

Feature matching using OpenCV is even faster than light-striping due to fewer required image processing steps and highly optimized search algorithms, in particular the Fast Library for Approximate Nearest Neighbor (FLANN).

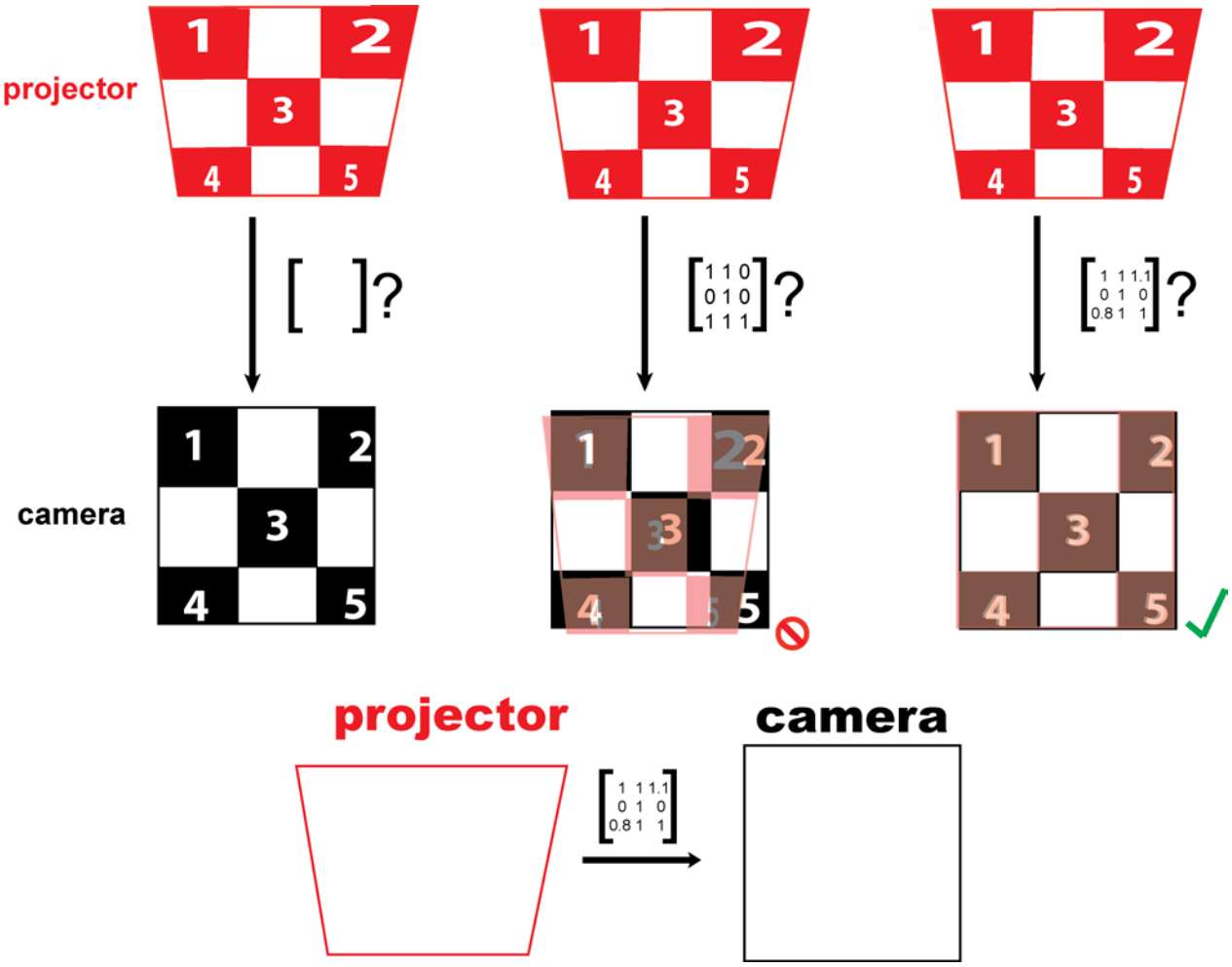


Figure 1-6. *Image registration by automatic feature matching. In this example, as in the previous figure, we are registering the projector image (red trapezoid) with the camera image (black square).A transformation matrix is required to transform the entire projector image into the camera image (left column). Features are automatically extracted, an initial transformation matrix is formulated and the resulting transformation compared to the camera image (middle column). The transformation matrix is iteratively optimized (right column). The optimal*

transformation matrix is recorded and used for all subsequent transformations (bottom)

Despite their strengths, both light-striping and feature matching performed poorly when the input images contained room dust. Dust particles are highly fluorescent, which caused light striping to generate ambiguous barcode, matching multiple projector pixels to a single camera pixel, or *vice versa*. Dust causes feature matching to detect non-existent features, generating non-sensical transformation matrices. By comparison, using control point matching, users can visually pick the crosses quickly and correctly regardless of minor target gel defects.

SI Gel Imager Acquisition of High Dynamic Range Images

With image registration in place, SILab acquires a series of 16-bit images of a gel where the exposure time for each successive image is doubled. All pixels within a given image that exceed a user-set threshold (default = 30,000 counts) are masked in all subsequent images of the SI series by turning the corresponding projector pixels to black, thus generating the mask. The mask is transformed to the projector coordinates as described above. The exposure time is doubled and another image is acquired with the mask in place. Since CCD cameras have highly linear responses up to saturation (Hiraoka et al., 1987) and exposure times are known, it is possible to calculate the expected intensity of a masked pixel by modeling over the intensity values of unmasked exposures. We set the masking intensity threshold to 30,000 to minimize camera noise found at the extremes of the camera's detection limits. To compile this series of masked images into a single high dynamic range image, a linear response curve was calculated, expressed in counts per second (CPS) for all unmasked pixels. SILab applies to each pixel a least-squares linear regression of the form

$$Y(c,t) = c_0 + c_1 t \qquad \text{(Equation 1)}$$

, where Y is observed intensity in counts, t is exposure time in seconds; c_0 and c_1 are the best-fit offset and slope. Using the GNU Scientific Library (Free Software Foundation, Boston, MA) mathematical toolkit, this interpolation is performed using the gsl_linear_fit() function which takes integer arrays for Y and t and explicitly returns c_0 and c_1. The coefficient c_1, the slope term of the linear least-squares fit, represents the intensity response of a pixel. Goodness-of-fit is calculated for each pixel

$$\%\varepsilon = \frac{\left(\sum_{i=1}^{n} \sqrt{\frac{(y-\hat{y})^2}{n}}\right) \times 100}{CPS} \qquad \text{(Equation 2)}$$

where %ε is the pixel's percent residual, y is the observed intensity, $\hat{y}$ is the interpolated intensity, n is the number of unsaturated exposures. The resultant SI image composed of C_1 coefficients for each pixel will be referred to as a CPS image.

To illustrate the CPS image calculation process, alcohol dehydrogenase (ADH) was labeled with Cy3-NHS dye using 10X the suggested concentration of minimal labeling dye to give Cy3-ADH, and 100 pg and 1 ng of this material was electrophoresed on a 12% SDS-PAGE gel (ADH, Sigma; Cy-dyes, GE Healthcare). A series of 5 cycles of masked SI images were taken with 0.125 s, 0.250 s, 0.5 s, 1 s and 2 s exposures (Fig. 1-2D, left side of the panel). Three individual pixels were chosen to highlight. Pixel A is in the brightest part of the 1 ng Cy3-ADH band. With the masking threshold set at 30,000, pixel A was masked in the 2 s exposure, therefore its intensity in the 2 s image was not used in the CPS calculation. Linear regression of the first four exposure values yielded a slope of 15,280 CPS. Dim pixels, such as pixel B, which contained no protein and pixel C, which was in the middle of the low-abundance Cy3-ADH band, were not masked since they never exceeded the masking threshold. All recorded intensity values for B and C were used to calculate CPS values of 526 and 6729, respectively. The first

exposure of an SI series was typically set to 0.1 s, ensuring that even highly abundant protein spots had at least three unmasked exposures from which to calculate CPS values. Each subsequent exposure time was double the previous, balancing obtaining maximum image data with a reasonable imaging time. The same linear regression was used for all pixels in the image, generating a single CPS image from the original five single-exposure images.

During the course of developing the CPS image, we sporadically encountered a visual artifact in test images, illustrated in Figure 1-7. The artifact, which we called “terracing”, consisted of a ring of darkened pixels around very bright protein spots, disrupting the normally smooth Gaussian contours of these spots (panel A). When we recovered the original pixel intensity values from the images that made up the CPS calculation, it became evident that these rings occur at the edge of the projected masks in the images. We chose three different pixels in the terraced image to illustrate the problem in panel B. Pixel A was within a bright area of the gel, and became masked after 3 exposures. Pixel A’s CPS value is correctly interpolated from these 3 unmasked exposures. Similarly, pixel C was in a darker area of the gel, and its 5 unmasked exposures resulted in a correct CPS value. In contrast, pixel B contained reliable intensity values for the 1s, 2s, and 4s exposures. However, in the 8s exposure several pixels surrounding pixel B were turned to black to generate the mask, and due to imperfect projector extinction, pixel B’s intensity was partially diminished. Interpolating over the intensity values from 1s, 2s, 4s and the darkened 8s exposure generated a lower CPS value for B than expected. This made B appear dimmer than expected in the CPS image, and pixels like similarly located at the edge of the masked area generated the darkened ring of pixel observed. These rings only appeared in images where there are highly-abundant spots that require several levels of masking, and therefore the terracing problem occurred only sporadically.

We employed a simple fix for the terracing problem. The CPS image is calculated at the end of the image acquisition process: SILab assembles a vector of exposure time values (t in Equation 1), a vector of intensity values (Y in Equation 1). To ensure only reliable intensity values are used in the CPS calculation, we added a check condition in the SILab source code to require that each intensity value is approximately twice the previous value, that is, $Y_{n+1} > 2*Y_n$. This check is given by (1) we're doubling exposure times during acquisition and (2) SI Gel Imager's CCD camera has very linear response. If the check is not met, both t and Y vectors are truncated to only include values that met the check. We found this check robustly fixed the terracing problem, as illustrated in panel C of Figure 1-7 using a 1D gel, while maintaining the accuracy of the CPS calculations.

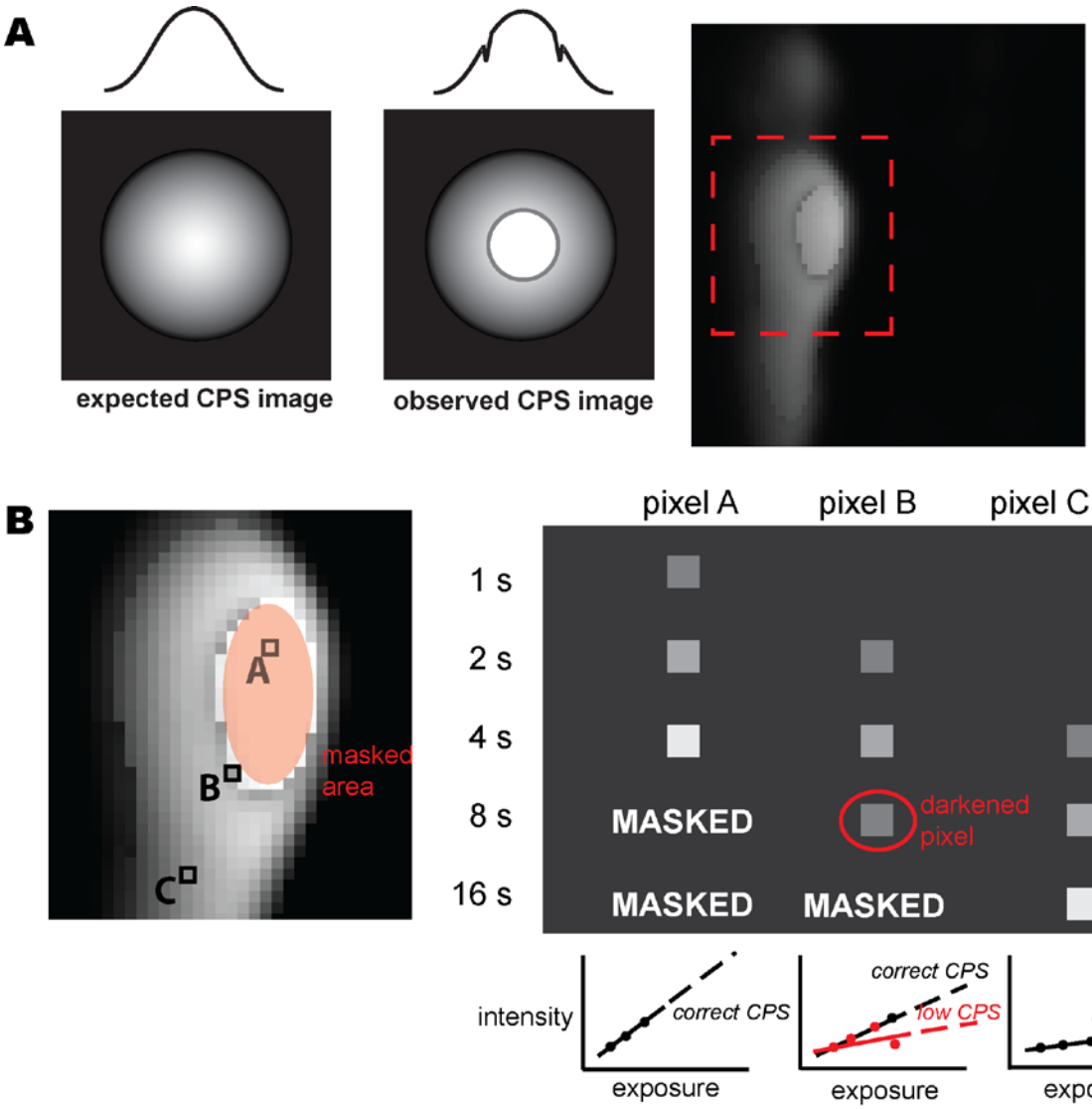

A
expected CPS image
observed CPS image
B
A
B
C
masked area
pixel A
pixel B
pixel C
1 s
2 s
4 s
8 s
16 s
MASKED
darkened pixel
MASKED
MASKED
intensity
correct CPS
exposure
correct CPS
low CPS
exposure
correct CPS
exposure
C
intensity
correct CPS
low CPS
exposure
"terraced" CPS image
brightest pixel excluded from CPS calculation
exposure
corrected CPS image

Figure 1-7. *"Terracing" artifact in early CPS images.* ***(A)*** *shows an idealized labeled protein spot showing the characteristic Gaussian shape (left). However, we observed that in some CPS images, some very bright spots showed a "terracing" appearance with a darker ring consisting of darkened pixels (center and right).* ***(B)*** *shows an enlarged version of the image in (A), which consists of intensity information from five exposures from 1s to 16s. Three pixels from this image were chosen to illustrate the problem. Pixel A and C are inside the masked area and their CPS values were correctly interpolated from reliable intensity values. The area surrounding pixel B was partially masked in the 4s exposure, leading pixel B to have* ***(C)*** *we corrected the terracing problem by excluding unreliable intensity values from the CPS calculation.*

In storing CPS image data we needed a data format that is (1) open, so we are not tied to any particular vendor and (2) well-documented and simple, so tools can be easily written to manipulate and display SI Gel Imager's high dynamic range images. To this end we considered OpenEXR (Kainz et al., 2004), an open-source image format developed for the motion-picture industry which handles 16-bit and 32-bit floating-point pixels. OpenEXR also had source code and reference viewer programs available. However, we find that OpenEXR's format to be unnecessarily complex for our purposes, having data structures for multiple image layers, required data conversions between RGB and luminance-only images, and in some cases, requiring specialized computer graphics cards for displaying images. Next we considered Flexible Image Transport System (FITS), a standard astronomy image format created by NASA's Goddard Space Flight Center (Pence et al., 2010). Designed for astronomical data, FITS supports extremely high dynamic ranges. Having been in near continuous development in the last 30 years meant FITS had a very large user community and robust tools for reading, writing and

manipulation. Unfortunately the format had also become too complex for our purposes, with many specified requirements particular to astrometry.

Instead, we devised our own simple file format for CPS images inspired by the file format of DeltaVision (DV) microscope image files (file extension .dv), which is readily handled by common image analysis software packages, such as ImageJ. Two .dv files are created in each imaging session, a temporary file containing the raw pixel intensity values from the camera, and the final CPS file containing the calculated CPS values. The dataflow of an imaging session is shown in Figure 1-8.

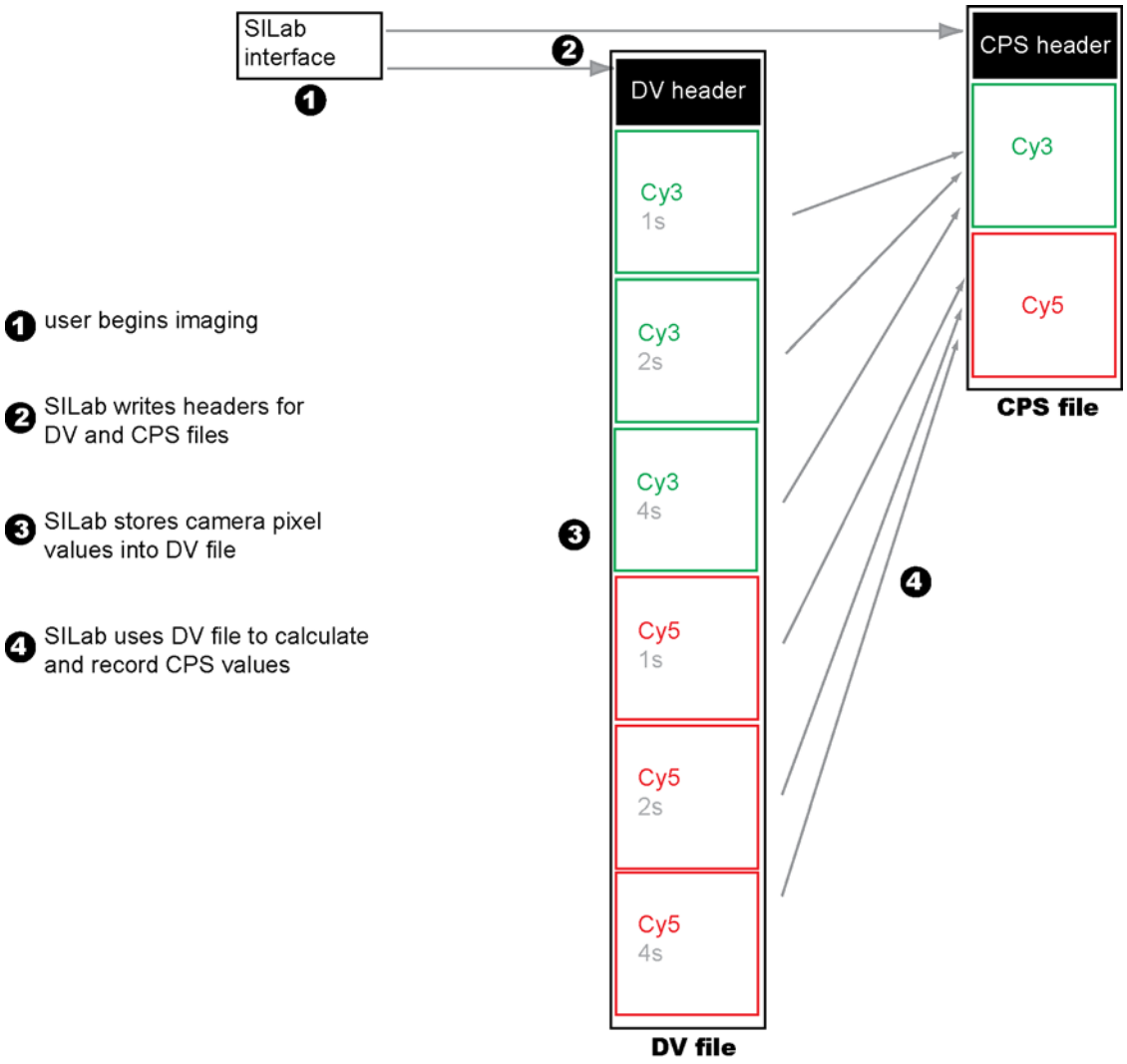


Figure 1-8. *Image data during a typical SILab imaging session. (1) The user selects the number of exposures desired and begins imaging. (2) SILab generates headers for the temporary .dv file and the CPS file. (3) Raw camera pixel values are recorded into the temporary .dv file as images are acquired (4) After acquisition is complete SILab computes CPS values and record them into the CPS file.*

The header of the CPS file is a 2048-byte segment, consisting of a mandatory first 1024-byte segment specifying the datatype of pixel intensity, image dimensions, number of image channels and an optional second 1024-byte segment. A non-exhaustive list of header descriptors is provided in Table 1-1. More detailed descriptions of the CPS file format's header and data sections are found in SILab's source code's `saveFileHDR()` and `writeCPSFile()` functions, respectively.

Byte offset	**Default value**	**Description**
0	null	image width in pixels
2	null	image height in pixels
4	null	number of images
6	6	datatype, 6 = 'unsigned short'
16	null	number of wavelengths
18	null	number of timepoints
46	1024	header size in bytes
99	null	excitation wavelength, in nm (used by some readers to false-color images)

Table 1-1. *Description of the CPS image file format header. This list is illustrative and therefore non-exhaustive.*

Post-capture Image Processing

Since complete camera coverage of a typical 2DE gel requires 4x5 tiles (1024 x 1280 pixels), SILab performs automatic stitching after image acquisition to generate a single image capturing the full gel. Previous work has determined a mathematical model for systematic variations in images of large format 2DE gels with an aim of eliminating them, particularly tiling artifacts (Sellers et al., 2007). When applied to SI images, this model appears as:

$$\pi_{x,y} = (Y_{x,y} - \rho_{x,y}) / \alpha_{x,y} \qquad \text{(Equation 3)}$$

, where π is the true abundance of proteins present in the gel, Y is the intensity of the observed image, ρ is the fluorescence emission of a blank polyacrylamide gel under uniform illumination, α is the fluorescence emission of a uniformly fluorescent (the bright-field image) gel target containing labeled proteins under the respective fluorescence filter, and x and y are pixel coordinates. To obtain ρ, we imaged an empty 12% polyacrylamide gel. For α, we prepared a uniformly fluorescent polyacrylamide gel: 1 µg of bovine serum albumin (BSA, New England Biolabs) was dissolved in the polyacrylamide gel solution before polymerization (detailed instructions for making this uniform gel is found in Appendix A2). To fluorescently tag the BSA within the gel with cysteine-reactive Cy3, the gel was then soaked in 40 mL of 1 mM tris(2-carboxyethyl)phosphine (TCEP) in a petri dish in the dark with shaking for 1 hour at 37 °C to reduce any disulfide bonds in the BSA, 4 µL of 10 mM Cy3-maleimide was then added and shaking continued for an additional hour. The uniform gel was then equilibrated in destain (40% methanol, 10% acetic acid) for 15 minutes and imaged. To minimize artifacts from fluorescent dust particles or local inhomogeneities in the gel, the median image from nine 1-second

exposures at slightly different XY positions relative to the camera field-of-view were used for the value of α. The image of true protein abundance π was then obtained computationally through ImageJ after image acquisition by solving the equation above.

Protein spot excision for sequencing by mass spectrometry

Somewhat tangential to the imager's gel imaging functions is its integrated robotic gel cutter. Mounted over the gel surface within the light-tight enclosure, the gel cutter is actuated by two linear motors to respectively lower and excise a 1mm-diameter gel plug, controlled by SILab's Cutter functions. These plugs can be deposited into 96-well plates and as of October 2013, 1.5mL tubes for trypsin digestion prior to sequencing by tandem MS.

Evaluating Imager Detection Capability

To determine the imager's practical dynamic range we prepared dilutions of purified protein samples in a typical SDS-PAGE mini-gel (which will be referred to as a 1D gel). We generated a 10-fold dilution series of Cy3- and Cy5-labeled ADH to create samples ranging from 10 µg to 10 pg of protein per lane. These samples were prepared from stock solutions where 240 µg of ADH was labeled with 10 nmol Cy3-NHS or Cy5-NHS dyes in 60 µl reactions, which represents a 10-fold increase in fluorescent labeling relative to the manufacturer's recommended minimal labeling. The stock solutions were diluted with an equal volume of 2X Laemmli sample buffer to a final concentration of 2 µg/µl. To eliminate excess unbound dye, the samples were then centrifuged in a Centricon 10 kDa spin dialysis device (EMD Millipore, Billerica, MA) at 4°C and 5900 rpm for 90 minutes. The 10-fold serial dilution series was made by mixing 10 µl of sample with 90 µl of diluent (100 µl of diluent contained: 47.5 µl of 2x Laemmli buffer, 2.5µl

βME, 10 µl of unlabeled 10 mg/ml BSA as carrier protein and 40 µl H_2O). BSA was added as a carrier to avoid protein loss due to adsorption, which is particularly important for the very low abundance samples. All samples were then boiled for 2 minutes and 5 µl of each sample was loaded into consecutive lanes of a 12% polyacrylamide mini-gel in increasing concentration to minimize cross contamination from adjacent wells. Electrophoresis was performed at 130 V until the tracking dye front ran off the gel, approximately 90 minutes. These 1D gels were removed from their electrophoresis plates, placed on the fused silica window of the gel tray containing a covering layer of destain (50% methanol, 10% acetic acid), and imaged for 13 exposures with masking threshold set at 30,000 counts and first exposure time set at 0.01 s. The resultant CPS images were analyzed in ImageJ: a rectangular selection box was manually drawn around the central ADH band in each lane and the pixels contained within quantified. The same selection box size was used for the lanes containing 10 pg to 1 µg, while the box for the 10 µg lane was increased in size to accommodate the high amount of protein loaded in this lane (Fig. 1-3).

To evaluate the SI gel imager performance with complex samples, we prepared large-format 2DE gels containing *Drosophila* embryo extract (DEE): 100 µg of DEE (obtained as described previously (Gong et al., 2004)) was labeled with 1 nmol of Cy3-NHS and incubated on ice for 20 minutes; 1µl of quencher (5 M methylamine-HCl and 10 mM pH 8.0 HEPES) was added and the reactions were then incubated for 2 hours before being run on 18 cm 3-10NL IEF strips–this labeling extent will be referred to as 1X-labeled (dyes and IEF strips: GE Healthcare, Uppsala, Sweden) for 32,000 volt-hours. Second-dimension SDS-PAGE was done in 18x20 cm gradient gels (10-15% polyacrylamide), electrophoresed in Tris-Glycine-SDS running buffer at 300 volts until the tracking dye ran off the bottom of the gel (5-6 hours). Gels were then removed from running cassettes, equilibrated overnight in destain (50% methanol, 10% acetic acid) to

proteins within the gel and imaged between fused silica plates the next day. For the 10X-labeled samples, 300 µg of DEE and 30 nmol of Cy3-NHS were used, the workflow was otherwise the same. Two images of each DEE gel were obtained: a 32-bit CPS image generated from 12 structured-illumination exposures with masking threshold set at 30,000 counts and first exposure time set at 0.1 s; and a single-exposure 16-bit TIFF image using a single 1 s full-field exposure. For quantitative analysis, the gels were cropped to the central 2x2 tiles (512 x 512 pixels) to eliminate artifacts such as streaks that commonly occur at the sides of the gels. Total image fluorescence (referred to as flux) was measured using ImageJ. Protein spots were detected and quantified using the DeCyder Differential Analysis 5.0 software suite (GE Healthcare), set to detect 400 spots in all gels. DeCyder employed a "watershed"-like detection method, segmenting the input image into patches containing pixels exceeding an internally-calculated background value and returning the centroid of each patch. The 2DE DEE images were also analyzed using SourceExtractor (abbreviated as "SExtractor" (Bertin and Arnouts, 1996)), an open-source astronomy software package that detects point sources in dim images that we had used to analyze 2DE gels previously (Minden, 2007). Conceived as a galaxy counting tool, SExtractor subtracts background signal from input images to look for Gaussian point sources that satisfied oval shape constraints. Though fundamentally different in their approaches, both DeCyder and SExtractor showed similar trends in results, as discussed later. Data analysis was done in Matlab (Mathworks, Natick, MA) and Excel (Microsoft, Redmond, WA).

RESULTS AND DISCUSSION

The experiments below were conducted to assess the gel imager's detection capabilities with contrived and real-world biological samples.

Million-fold Detection of Purified Protein Samples in 1D gels

To assess the detection range of the SI gel imager, a CPS image of a 1D gel containing the ADH million-fold dilution series was acquired (Fig. 1-9). The CPS image was computed from a series of 13 progressively masked images where the exposures ranged from 0.01 to 40.96 s, and the masking threshold was set at 30,000 counts. Protein labeling was done with Cy3-NHS dye at ten times the usual, commercially recommended dye-to-protein concentration (referred to as 10X-labeling). Progressively decreasing the image's displayed maximum pixel intensity value allows one to see the dim protein bands (Fig. 1-9, 1^{st}-5^{th} rows, minimal and maximal display values are shown to the left of each row). This high dynamic range CPS image shows that the entire million-fold concentration range was detectable in a single image, which is not possible with conventional 16-bit images. To visualize the full concentration range, the image is displayed using a natural logarithmic scale look-up table (Fig. 1-9, 6^{th} row). Horizontal line plots show that the samples form distinct bands of decreasing intensity from the lane containing 10 μg of protein down to the lane containing 10 pg (inset images). The CPS image had a maximum intensity of 2.13 million counts, far exceeding the 65536 count upper limit allowed by single-exposure 16-bit images. The integrated intensity of the 10 pg band was 5.12×10^4 CPS.

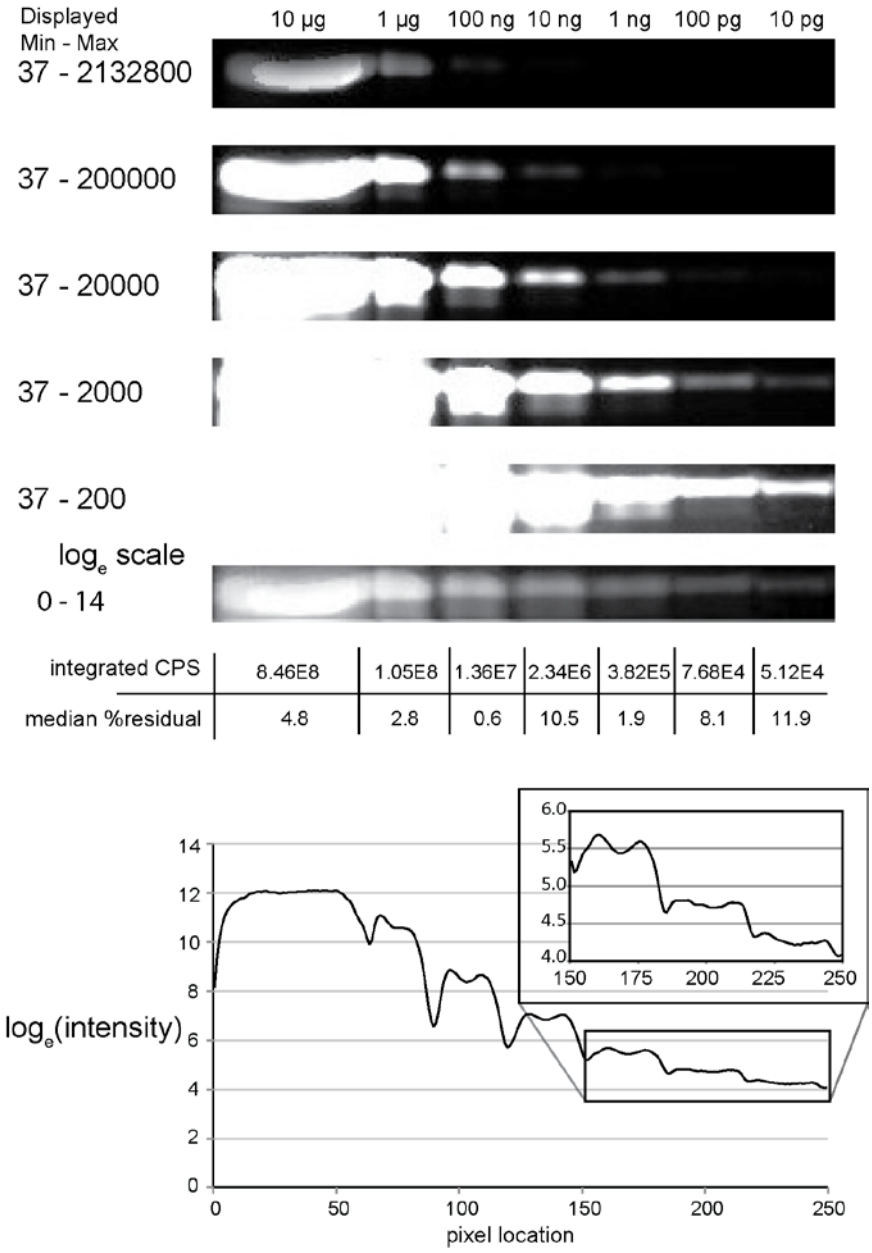


Figure 1-9. *The SI gel imager detects protein over a 1,000,000-fold concentration range. Seven 10-fold dilutions of purified, Cy3-ADH from 10 µg to 10 pg were electrophoresed in a 1D SDS-PAGE gel and imaged with the SI gel imager to generate a CPS image. Successively decreasing the displayed maximum pixel intensity reveals progressively dimmer bands (1st-5th row). All seven bands, covering a million-fold concentration range, are visible in the same CPS image when the image is displayed with a logarithmic (log$_e$) look-up-table. Horizontal line plots of this*

log-transformed image show distinct bands of decreasing intensity from 10 mg to 10 pg.

The calculated median residual of the linear fit of this area, which represents the measurement error, was found to be 11.9%. A background region having the same area as the 10 pg band had an integrated intensity value of 2.62×10^4 CPS, with a median residual of 1.9%. Thus, the 10 pg band fluorescence was significantly brighter than the gel's background fluorescence.

It is important to note that this result was only possible with 10X-labeling. The same million-fold dilution experiment with ADH labeled at the recommended Cy3-NHS dye concentration (1X-labeling) only provided a detection range of the five most abundant lanes in the CPS image, a ~10,000-fold dynamic range (Fig. 1-10).

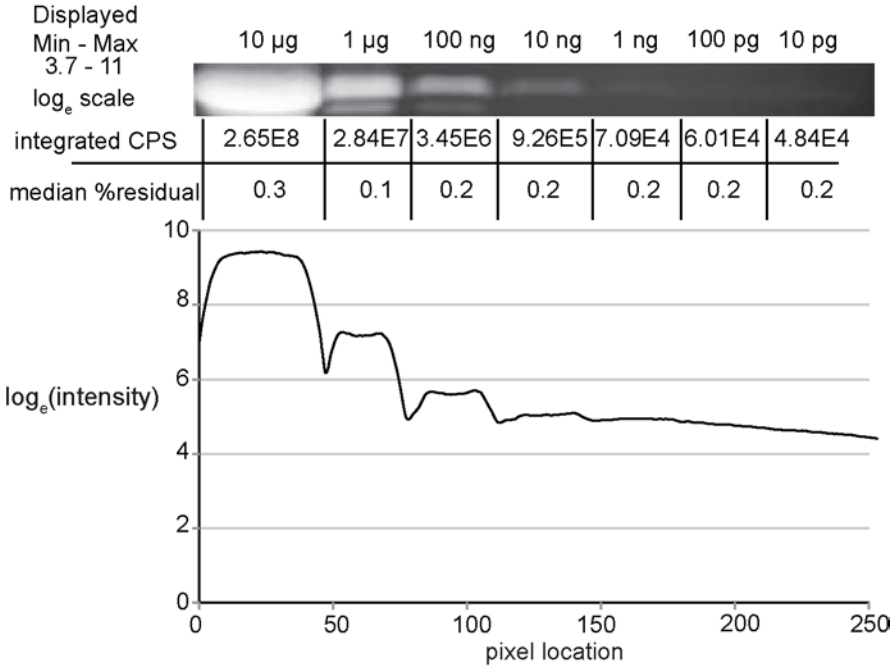


Figure 1-10. *Million-fold dilution series (10 µg - 10 pg) of 1X-labeled Cy3-ADH. Only the most abundant 5 lanes (10 µg to 1 ng) are visible in the CPS image.*

Similarly high dynamic range images were captured for Cy5-NHS-labeled ADH,

indicating the SI gel imager performance is independent of the fluorophore used, allowing for the detection of low-abundance protein differences using DIGE labeling schemes (Figure 1-11).

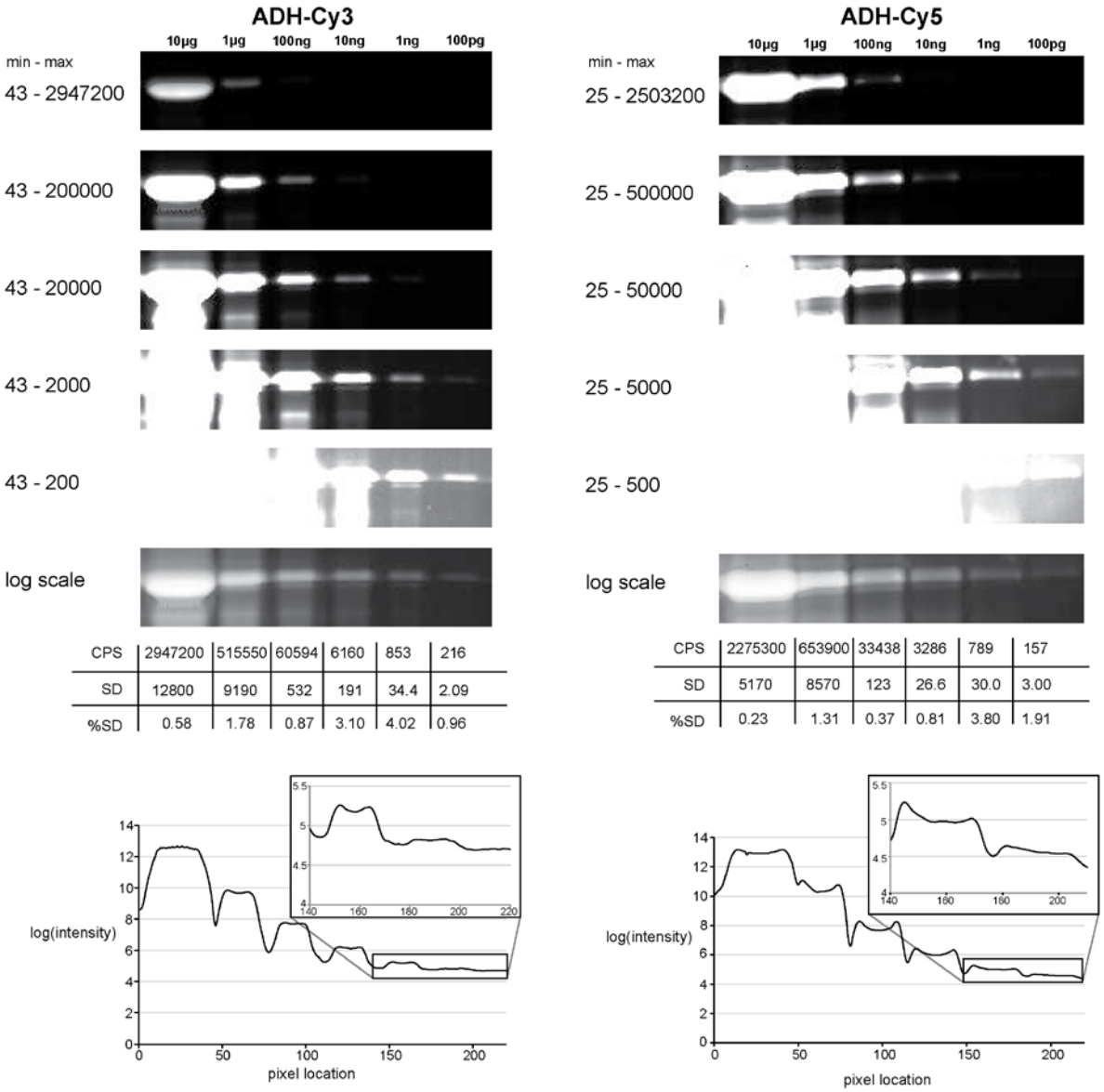


Figure 1-11. *The SI gel imager performs equally well with Cy3 and Cy5. An early paired 100,000-fold dilution series using ADH labeled with Cy3-NHS and Cy5-NHS showed similar performance, leading us to perform subsequent experiments using only Cy3 for the sake of simplicity.*

The aforementioned million-fold concentration range experiment was challenging to perform because minute amounts of Cy3-ADH from the highest abundance lane tended to leak into the upper tank buffer of the gel apparatus, thus contaminating all lanes of the gel. Loading the highest abundance lane last, as well as spin dialyzing the samples to remove unincorporated dye, mostly alleviated this contamination problem, however there was persistently a slight remnant of fluorescent Cy3-ADH observed in blank lanes of the gel. While the lowest abundance lanes were significantly detectable above the background fluorescence, this slight contamination made quantification of the low abundance difficult. To confirm the detection and quantification of the 10 pg band, we ran a thousand-fold dilution series on a 1D gel using the lowest four spin-dialyzed samples (10 ng to 10 pg). In the resulting CPS image, the 10 pg band is clearly visible above background. There was no detectable fluorescence signal in lane adjacent to the 10 pg lane (Fig. 1-12). The residual for the 10 pg band, which had an integrated fluorescence signal of $2.67x10^4$ CPS, was 7.6%, making it distinguishable from the integrated intensity of $1.97x10^4$ CPS for the neighboring empty lane, which had a residual of 8.8%.

Taken together, these experiments demonstrate that the SI gel imager has an effective dynamic range of at least 1,000,000-fold in 1D gels loaded with 10X-labeled proteins.

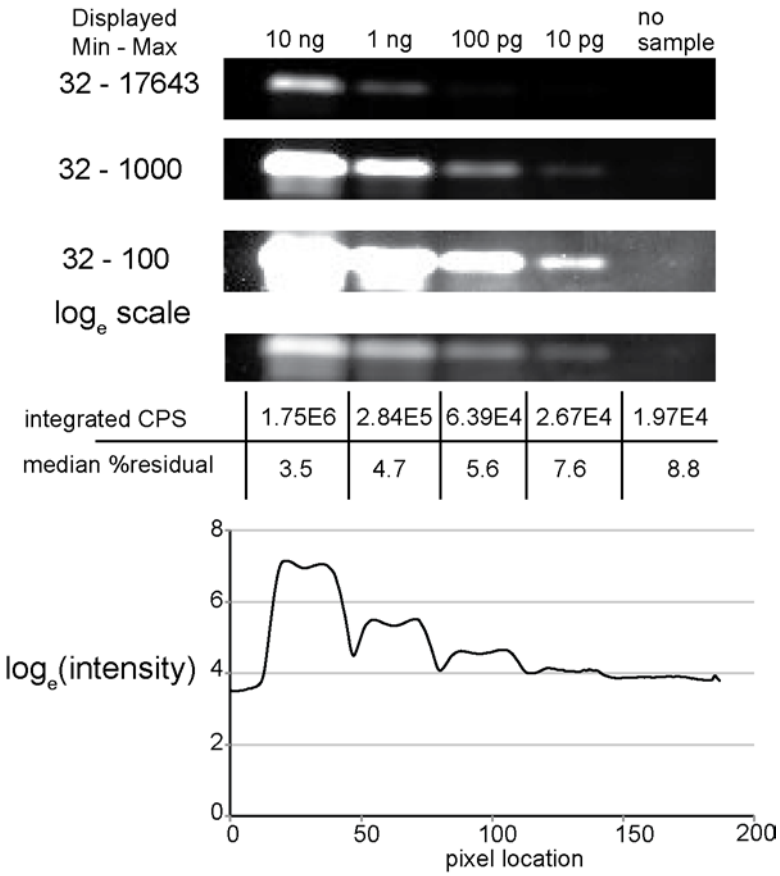


Figure 1-12. *Thousand-fold dilution series (10 ng to 10 pg) of ADH labeled with 1X Cy3-NHS. The last four samples of the spin-dialyzed, 10X-labeled Cy3-ADH dilution series from Fig. 1-9 were imaged with the SI gel imager. This image clearly shows the presence of the 10 pg sample relative to the adjacent empty lane.*

Masking Improves Local Image Contrast Under Certain Conditions

During the development of the SI Gel Imager prototype, we observed that the structured illumination scheme improved the local image contrast when a highly abundant protein spot is located next to a low-abundance spot. The decreased illumination on the bright spot allowed the dim spot to be more easily distinguished when compared to full-field illumination. This was demonstrated using BSA labeled with Cy3-maleimide, where 10 μg of Cy3-BSA was loaded

next to a lane containing 10 pg of Cy3-BSA in a 1D polyacrylamide gel (Fig. 1-13).

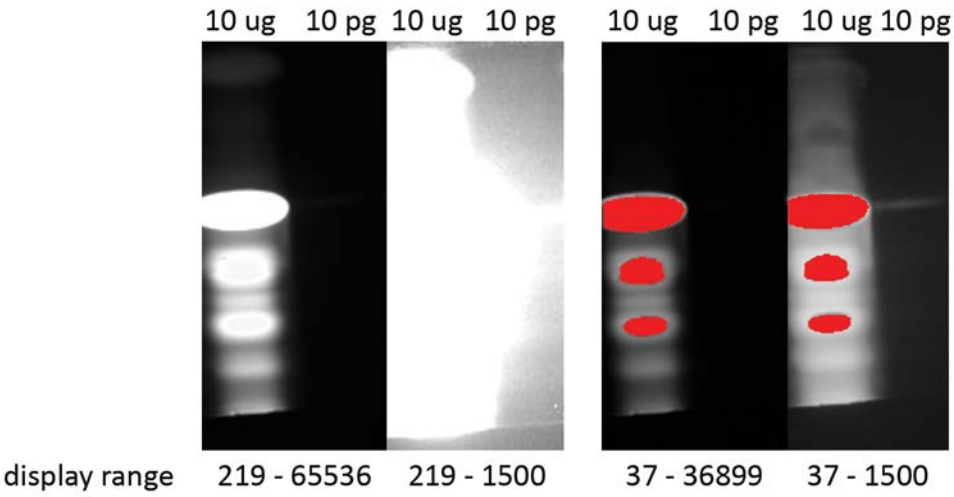


Figure 1-13. *Masking improves local image contrast under certain conditions. In an early prototype of SI Gel Imager, we observed that a 10 pg sample of Cy3-BSA can be readily seen next to a 10 μg sample even if the two were loaded in adjacent wells in a 1D polyacrylamide gel (right panel, masks false-colored red for clarity). The image pair on the left is single shot at full intensity range and scaled to detect the 10 pg band. The image pair on the right is a CPS image scaled to full range and re-scale to observe the 10 pg band.*

However, we were unable to replicate these results in the production imager, leading us to believe that this improvement in local contrast only existed due to the relatively poorer light seals of the prototype imager. Furthermore, even if we had observed this improvement in the production imager, the performance gain is small and difficult to predict and quantify in 2DE gels using real-world samples.

<u>High Dynamic Range Detection of Complex Protein Mixtures in 2D gels</u>

To assess the efficacy of the SI gel imager for the analysis of cellular proteomes, a series

of 2DE gels containing Cy3-NHS 1X-labeled *Drosophila* embryo extract was analyzed. The goal was to compare the detection range and reliability of single-exposure (un-masked) images to SI gel imager-generated CPS images. These images were analyzed using several image analysis packages: (1) DeCyder (version 5.0), a commercial, 2DE gel analysis software package from Amersham BioSciences (now GE Healthcare Life Sciences), (2) Source Extractor (SExtractor, version 2.8), an open-source astronomy analysis application that was adapted to analyze 2DE gel images and (3) ImageJ (version 1.39j), a well known public, biological image analysis package from the US National Institutes of Health. The integrated fluorescence intensity of a region of the image, which could encompass a single protein spot, a set of protein spots, or large swaths of an image, will be referred to as the flux (borrowing the term from astronomy). Each image was processed to account for background fluorescence (dark-field correction) and to correct for variation in projector output across the imager field of view (bright-field correction). These correction elements yield images where the pixel values closely represent the fluorescent signal emanating from labeled protein, thus limiting artifacts arising from the imaging system.

Analysis of the 2DE gel images with DeCyder revealed advantages of CPS images over single-exposure images using the same DeCyder detection parameters. Using standard parameters to assess proteins in the central 2 x 2 tiled region of a full 2DE gel (Fig. 1-14A), DeCyder detected an 8% (336/311) increase in raw spot count between the single-exposure image and the CPS image (Table 1-3). Since 2DE gels of fluorescent proteins tend to be noisy, unsupervised, raw spot counts by DeCyder tend to contain a significant fraction of falsely detected protein spots. All detected spots in the single-exposure and CPS images were visually inspected to validate the DeCyder spot detection.

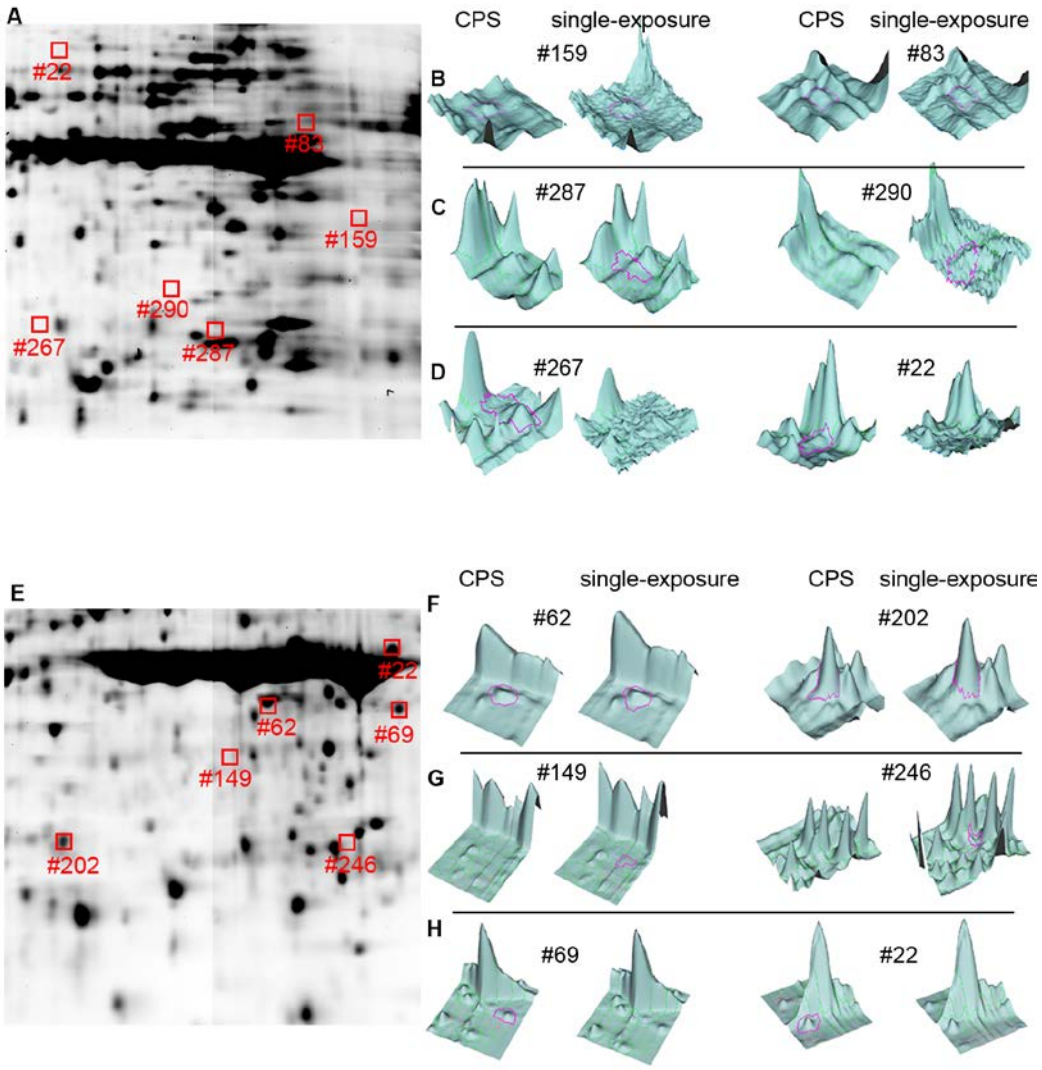


Figure 1-14. *Comparison of DeCyder spot detection of CPS and single-exposure images.* ***(A)*** *Shows a CPS image of a 1X-labeled Cy3-DEE sample resolved on a 2DE gel. DeCyder was used to detect protein spots in CPS and single-exposure images of this 2DE gel. The numbered protein spots indicate a set particular spots chosen to highlight the differences between the CPS and single-exposure images.* ***(B)*** *Shows surface plots of two protein spots positively detected by DeCyder in the CPS and single-exposure images of the gel shown in A. The magenta boundaries indicate the DeCyder detected protein spots.* ***(C)*** *Shows a pair of protein spots that were erroneously detected by DeCyder in the single-exposure image, but not registered by DeCyder in*

the CPS image of the gel in A. ***(D)*** *Shows a pair of protein spots that were correctly detected by DeCyder in the CPS image, but not registered by DeCyder in the single-exposure image of the gel in A.* ***(E)*** *Shows a CPS image of a 10X-labeled Cy3-DEE sample resolved on a 2DE gel.* ***(F)*** *Shows a pair of positively detected protein spots in the CPS and single-exposure images of the gel shown in E.* ***(G)*** *Shows a pair of protein spots that were erroneously detected in the single-exposure image, but not in the CPS image of the gel in E.* ***(H)*** *Shows a pair of protein spots that were correctly detected in the CPS image, but not in the single-exposure image of the gel in E.*

Visual inspection in the single-exposure image rejected 102 detected objects, reducing the number of detected spots by 33%. Visual validation of the CPS image leads the rejection of 53 putative protein spots, a 16% reduction (Table 1-2). Thus, about 50% fewer spots needed to be eliminated when analyzing the CPS image versus the single-exposure image. This indicates that the CPS image was a more reliable input for DeCyder spot detection. In terms of visually verified spots, the CPS image contained 35.4% more detectable spots than the single-exposure image of the same 2DE gel, showing higher image quality for a more complete characterization of the protein sample (Table 1-2).

To further illustrate the image quality differences between single-exposure and CPS images, the DeCyder's 3D SpotView was used to display surface plots of spot features. This perspective allows one to visualize real spots, as well as erroneously detected, in single-exposure versus CPS images (Fig. 1-14 B-D). Since the CPS image is a composite of multiple exposures, the resultant image is an averaged representation of the gel. The CPS image had lower noise in all parts of the gel, evident in its smoother contour plots. In some examples of dim spots, DeCyder was able to detect spots in the single-exposure image despite its higher noise (Fig. 1-

14B). However, image noise caused DeCyder to falsely assign spots to the "shoulder" of an existing spot (Fig. 1-14C, left panel), or to background areas (Fig. 1-14C, right panel). Conversely, the same DeCyder analysis revealed genuine spots detected in the CPS image that were overwhelmed by noise and were consequently not detected in the single-exposure image (Fig. 1-14D, both panels).

An important measure of the efficacy of the SI gel imager is to determine the dynamic range of cell extracts separated on 2DE gels. This was done by comparing the flux of the dimmest, verified protein spot relative to the flux of the brightest protein spot. The flux ratio for the single-exposure image was nearly 2,400, while the flux ratio for the CPS image was 12,000 (Table 1-2).

	Single-exposure image	CPS image
Number of detected protein spots	311	336
Number of verified protein spots	209	283
Dimmest protein spot (flux)	12,100	2,360
Brightest protein spot (flux)	27,800,000	26,900,000
Summed Yolk region (flux)	102,000,000	90,100,000
Flux ratio (Brightest / Dimmest)	2,300	11,400
Dynamic range (Summed Yolk / Dimmest)	8,500	38,100

Table 1-2. *DeCyder spot detection and quantification of a 2DE gel containing 1X-labeled DEE.*

Therefore, the dynamic range of 1X-labeled *Drosophila* embryo extract according to the CPS image was five-fold greater than the single-exposure image. The most abundant protein in *Drosophila* embryos is the Yolk Protein. There are three closely related Yolk Proteins that share several common peptide sequences, each of which migrates as several post-translationally

modified isoforms. The Yolk Proteins appear on 2DE gels as large series of very high abundance proteins. A previous study characterized the *Drosophila melanogaster* proteome using a LC MS/MS "shotgun proteomics" approach and reported a dynamic range of approximately 10,000-fold (Schrimpf et al., 2009). The high degree of Yolk Protein homology, which is 47-52% identity, means that these proteins share common tryptic peptides and that the different isoforms are indistinguishable by mass spectrometry. To compare SI imager dynamic range with the LC MS/MS results, we determined the maximum protein concentration by summing DeCyder-reported fluorescence volumes for Yolk Protein spots in the gel images (Yolk Proteins were identified on the gel image using their distinctive distribution pattern, which was previously confirmed by MS (Gong et al., 2004)). Comparing the flux of dimmest protein spot to the summed Yolk Protein signal increased the dynamic range to 8,480 and 38,121 for the single-exposure and CPS images, respectively. This represents a nearly four-fold increase in dynamic range. It also indicates a four-fold improvement in detected concentration range in *Drosophila* embryos.

The Effect of Excess Fluorescent Labeling on Protein Spot Detection

In the preceding section, we demonstrated that 1X-labeled, purified proteins separated on 1DE gels only displayed a 10,000-fold dynamic range using CPS images, which is roughly consistent with the dynamic range observed in the 2DE gels of 1X-labeled *Drosophila* embryo extract. In order to achieve the desired million-fold detection range, the extent of labeling required a 10-fold increase in labeling. 10X-labeled *Drosophila* embryo extracts were separated by 2DE to determine if the SI gel imager was able to detect an even greater concentration range.

	Single-exposure image	CPS image
Number of detected protein spots	326	318
Number of verified protein spots	279	293
Dimmest protein spot (flux)	10,400	18,100
Brightest protein spot (flux)	28,400,000	42,500,000
Summed Yolk Protein region (flux)	159,000,000	247,000,000
Flux ratio (Brightest / Dimmest)	2,750	2,350
Dynamic range (Summed Yolk / Dimmest)	15,420	13,700

Table 1-3. *DeCyder spot detection and quantification of a 2DE gel containing 10X-labeled DEE.*

DeCyder spot extraction for a 10X-labeled 2DE gel showed a similar spot density and a modest 5% increase in verified spot count comparing the single-exposure to the CPS image (Table 1-3). Increased labeling generated brighter fluorescent signal as expected, and DeCyder quantified more flux. For the single-exposure image, the dynamic range (flux ratio dimmest/brightest) for the 10X-labeled sample was similar to that of the 1X-labeled sample. Surprisingly, the dynamic range of the CPS image for the 10X-labeled sample was slightly lower than the single-exposure image of the same gel with the 10X-labeled sample. The same trend was also observed for the flux ratio of the summed yolk spots relative to the dimmest spot. We expected to observe an increase in dynamic range when comparing 1X- to 10X-labeled samples, based on the assumption that increased labeling would reveal more low abundance proteins. To investigate this result, we examined the Yolk Protein area of the 10X-labeled DEE gel. Despite the increase in fluorescence signal, the DeCyder measured flux emanating from the Yolk Protein

spots did not increase 10-fold, as expected. Instead, when compared to the same area in 1X-labeled DEE 2DE gels, the 10X-labeled Yolk Proteins produced spots that were larger in cross-section (full width at half-maximum intensity), but not significantly higher in peak intensity. The 10X-labeled Yolk Proteins spots appear larger mostly in the second dimension, not the first dimension, because of the increased labeling which reveals that these high-abundance proteins experience molecular crowding in the second dimension; they are forced to spread out further in this high percentage polyacrylamide gel because of molecular crowding in the gel matrix (Fig. 1-13). This molecular crowding may explain why the 10X-labeled Yolk Protein peaks did not have the expected Gaussian shape, but appeared to be somewhat plateaued (Fig. 1-15).

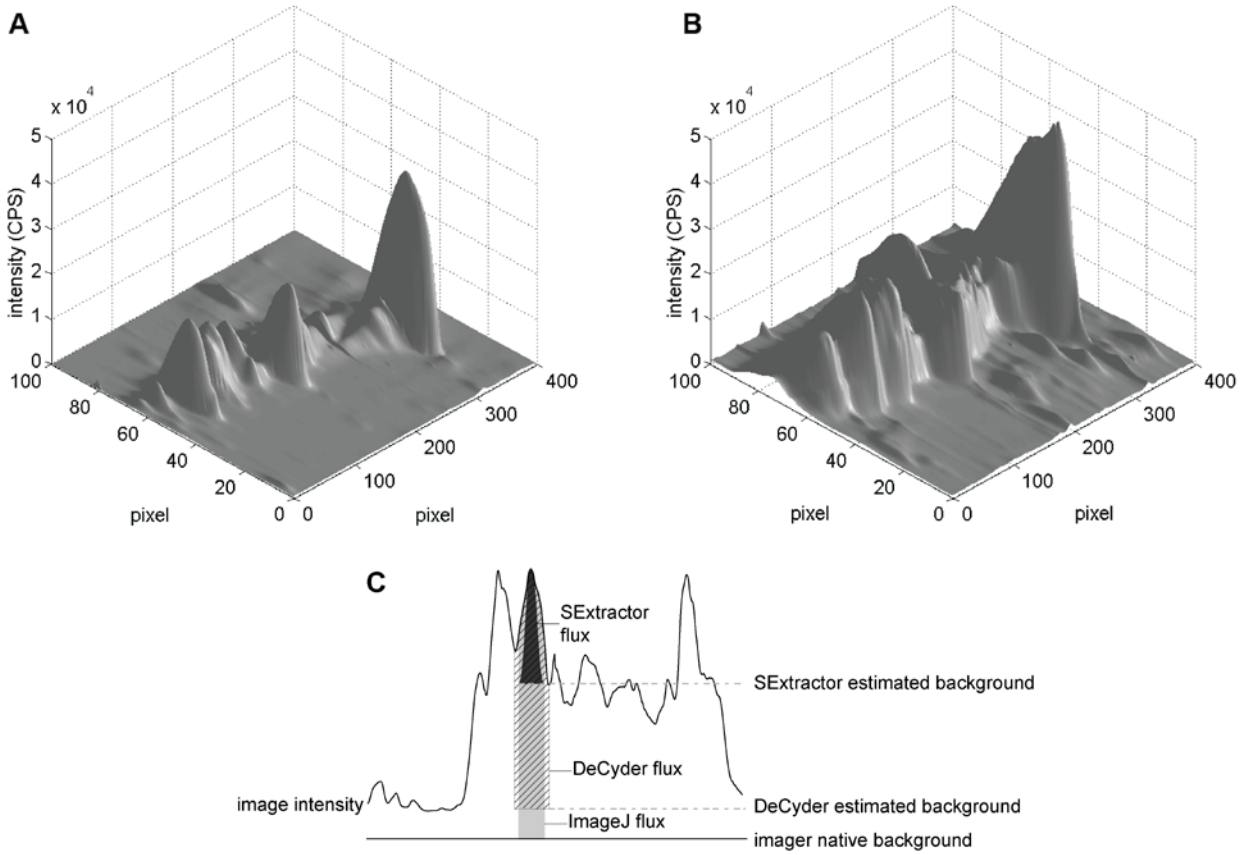


Figure 1-15. *Surface plots of the Yolk Protein regions of DEE 2DE gels.* ***(A)*** *Shows a Matlab-generated surface plot from the CPS image of the Yolk Protein region of a 1X-labeled DEE 2DE*

gel. ***(B)*** *Shows the same region of a 10X-labeled DEE sample.* ***(C)*** *Shows intensity cross-section of the Yolk Protein region of a DEE gel. To illustrate differences in flux estimation for a single protein spot, the area shaded black shows the SExtractor flux estimate with local background subtraction, the hashed area shows the DeCyder flux estimate with the background relative to the segment boundary subtracted, and the area shaded gray shows the ImageJ estimated flux based on the SExtractor aperture boundary and no additional background subtraction*

High-abundance Protein Flux Measurement Errors

The discrepancy in flux ratio between 1X-labeled and 10X-labeled protein samples prompted us to examine how DeCyder quantified protein spots. Comparing the flux of the dimmest spot detected in the CPS image of the 1X-labeled sample to the dimmest spot detected in the 10X-labeled sample, showed an approximate 10-fold increase in flux, as expected. It is important to note that these were not the identical protein spots in the gels being compared, they were simply the lowest abundant spots detected (Table 1-2 versus Table 1-3). Comparing the brightest individual spot and the summed yolk region of the 1X-labeled and 10X-labeled samples revealed only a 1.6- to 2.7-fold flux increase, respectively (Table 1-2 versus Table 1-3). While low abundance, well-separated protein spots roughly reflected the 10-fold increase in labeling extent, high abundance proteins failed to reflect the expected ten-fold increase in labeling. In DeCyder, the background value for any spot segment is taken as the lowest 10^{th}-percentile value of the spot border, not the gel's native background signal (Amersham, 2002). This approximation works well for isolated dim spots, where the segment border is very close to the background signal. However, for bright protein spots in crowded regions of the gel, as is the case the Yolk Proteins, the spot background value is significantly higher than the native background value.

Thus the height of the protein spot is truncated, eliminating pixel counts between the segment border down to the image background, which can grossly underestimate the total flux from highly abundant protein spots.

To circumvent DeCyder's potential underestimation of high abundance proteins, the images were also analyzed by SExtractor, an astronomical image analysis package. Overall, the results from SExtractor were similar to DeCyder in terms of detected spots and flux ratios (Tables 1-4 and 1-5). However, SExtractor also suffered from an underestimation of high abundance proteins. SExtractor uses a mixed Gaussian fitting algorithm to detect celestial objects, which we adapted for 2DE gels (Bertin and Arnouts, 1996) . There are many different, user-controlled parameters that allow for spot-detection optimization, including spot size range, spot packing and background estimation.

	Single-exposure image	CPS image
Number of verified protein spots	272	301
Dimmest protein spot (flux)	1570	704
Brightest protein spot (flux)	1,460,000	2,670,000
Summed Yolk spots (flux)	29,100,000	55,200,000
Flux ratio (Brightest / Dimmest)	932	3,790
Dynamic range (Summed Yolk spots/ Dimmest)	18,500	78,400

Table 1-4. *SExtractor/ImageJ spot detection and quantification of a 2DE gel containing 1X-labeled DEE.*

	Single-exposure image	CPS image
Number of verified protein spots	293	326

Dimmest protein spot (flux)	1400	2380
Brightest protein spot (flux)	1,310,000	3,500,000
Summed Yolk spots (flux)	158,000,000	169,000,000
Flux ratio (Brightest / Dimmest)	942	1,470
Dynamic range (Summed Yolk spots/ Dimmest)	113,000	71,100

Table 1-5. *SExtractor/ImageJ spot detection and quantification of a 2DE gel containing 10X-labeled DEE.*

However, one of the SExtractor's attributes that cannot be easily modified is the scope of background estimation. Inherent in the software is the assumption that all celestial objects are smaller than a pre-set size. Any detected object greater than this maximum object size is considered as a background source and its intensity contribution is removed. SExtractor produces a background image that is subtracted from the input image to yield a background-corrected image. The area occupied by very high abundance proteins exceeds the maximum object size limit of SExtractor such that part of the signal from very abundant proteins is considered background. Thus, SExtractor also underestimates the flux emanating from very high abundance proteins. In contrast to DeCyder which attempts to assign as much of the input image's area to spots as possible, SExtractor only quantifies areas that exceed its background threshold, further diminishing its dynamic range.

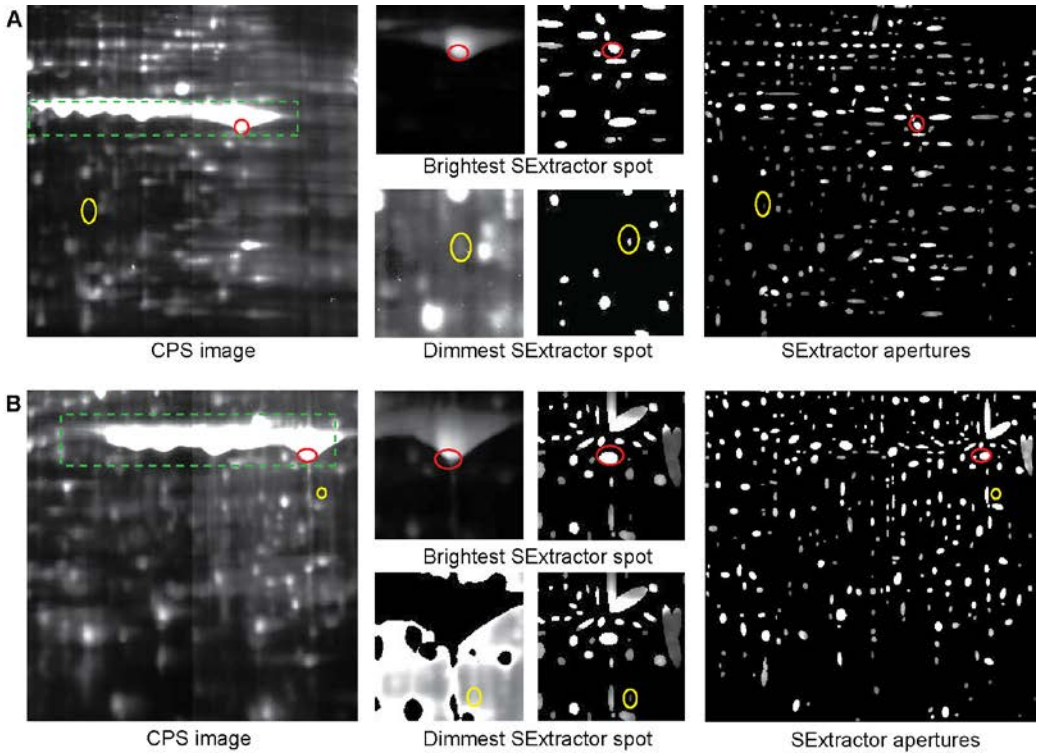


Figure 1-16. *SExtractor analysis of CPS images. The same DEE 2DE gels analyzed by DeCyder in Fig. 1-14 were analyzed using SExtractor. (A) Shows the 1X-labeled DEE 2DE gel. The left-most panel displays the CPS image, while the right-most panel shows SExtractor's detected apertures. The upper pair of smaller middle panels magnify the brightest spot of the image (indicated by the red oval), with the CPS image on the left and the SExtractor apertures on the right. The lower pair of smaller middle panels magnify the dimmest spot of the image (indicated by the yellow oval), with the CPS image on the left and the SExtractor apertures on the right. (B) Shows the 10X-labeled DEE 2DE gel. The panels are arranged similarly to A. The magnified image of the dimmest spot was masked with a threshold of 45,000 CPS to enhance the visualization of this very dim protein spot.*

Examples are illustrated in Fig. 1-16: the brightest SExtractor spot comprised only a small fraction of the yolk spot (Fig. 1-16A and B, the upper middle panels). The aperture

highlighted by the red oval shows that only a small portion of the high-abundance Yolk Protein spot contributed to the flux measurement for this protein. This was the case for both 1X- and 10X-labeled samples (Fig. 1-16A and B). SExtractor also failed to detect very low abundance proteins as evidenced by the visual detection of many spots in the vicinity of the dimmest, SExtractor-detected spot that are less bright (Fig. 1-16A, lower middle panels–highlighted by the yellow oval). SExtractor was capable of detecting very-low abundance protein spot near very high abundance protein spots (Fig. 1-16B, lower middle panels– highlighted by the yellow oval). The CPS image was masked with a threshold of 45,000 CPS to display this very dim spot.

Since SExtractor and DeCyder both underestimated the images' dynamic ranges, and other commercial 2DE spot analysis software packages suffer from similarly poor automated spot detection and quantification (Kang et al., 2009b), ImageJ was used to manually draw a rectangular selection box around the Yolk Protein area (Fig. 1-16, green, dashed-line rectangles). Since all SI gel imager images are corrected for both background and bright-field variation, there is no need to reject any portion of a pixel's intensity after these imager-dependent corrections are applied. SExtractor generates a list of elliptical apertures, where each aperture represents a detected object. Each aperture is defined by its location, the length of the major and minor axes and the angle of the major axis. To estimate the total flux of an aperture, the aperture was superimposed on the corrected CPS image in ImageJ and the total pixel count was determined for all pixels within the aperture. Thus, the measured flux from a spot is the product of SExtractor detection and ImageJ quantification (shown as SExtractor/ImageJ flux in Tables 1-4 and 1-5). The vast majority of protein within the Yolk Protein box is known to be Yolk Protein, therefore it is reasonable to perform a simple integration of the detected spots this region to estimate the total Yolk Protein flux using ImageJ. This approach yielded flux ratios for CPS

images of 78,370 and 71,086 for 1X-labeled and 10X-labeled samples, respectively (Tables 1-4 and 1-5). These data demonstrate that allowing the image analysis tool to rely on systematically corrected images, without further background correction, produced very high dynamic range estimates that are roughly equal between the 1X-labeled and 10X-labeled samples. These data also suggest that increased labeling does not increase the detectable range of a sample–the SI imager is capable of detecting the full protein concentration range with the *Drosophila* embryo extract resolved on a 2DE gel. In other words, if SI imaging of 1X-labeled sample was not sufficient to detect the full concentration range of proteins in a sample, then 10X-labeling would be expected to reveal a wider concentration range. Since the detected protein concentration range was roughly similar between 1X- and 10X-labeling, one can assume the SI gel imager detected the full range of proteins in these embryo extracts.

Excision of protein spots from 2DE gels for mass spectrometric analysis

After obtaining high-quality 2DE gels, a common task is to excise protein spots for peptide fingerprinting using MALDI-TOF or sequencing by LCMS/MS. This can be done by hand using a visible counter-stain like silver or Coomassie Brilliant Blue and cutting the spots with a razor, or using a commercial computer-controlled gel cutting robot. Previously the Minden lab had developed an integrated gel-cutter into the imager. Our typical workflow involves running 2DE gels over the course of 3 days (a more detailed schedule is presented in Chapter 3), image with SI Gel Imager immediately upon completion of the 2nd dimension and equilibrate gels in airtight lighttight containers overnight. The next day, final images are obtained and spots of interest are cut in a single session to minimize protein loss. Though we have documented some protein loss in destain, as shown in Figure 1-17, we find this workflow is the

best compromise of current technology and scheduling.

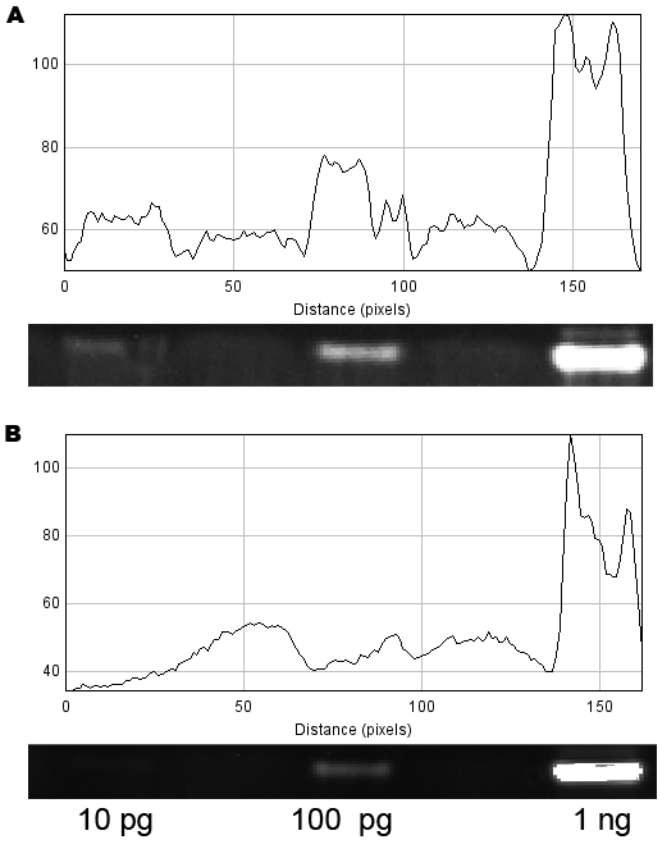


Figure 1-17. *Protein diffusion loss from polyacrylamide gel in destain. Samples of ADH labeled with 10x Cy3-NHS were electrophoresed in a 12% polyacrylamide gel before being soaked in methanol-acetic acid destain for 5 hours (A) and (B) 24 hours. Images are scaled to the same display range. Line plot represents average intensity of the gel areas shown.*

Part of my work also included improving the protein spot cutting routine. To this end, I have documented the major problems typically encountered by users of the SI gel imager, outlined in Table 1-6 and discussed in more detail in Appendix A-3

Problem	**Cause**	**Solution**

Gel plugs are dropped back onto gel plate surface	Inner plastic connecting rod is worn from repeated actuations, cutting head loses traction	Replace rod
Gel plugs are partially cut and still attached to gel ("chads")	Cutting cycle delay is too short, cutting head does not generate enough suction Cutting parameters are not optimized for gradient gels	Increase cutting cycle delay in SILab code Change cutting parameters when cutting in new areas of gel
Gel plugs are sheared or shattered	Aluminum sleeve of cutting head is deformed from repeated hard contact with fused silica plate	Reshape cutting head sleeve to restore circularity of cutting aperture, reduce vertical cutting travel to just enough to cut through gel

Table 1-6. *Troubleshooting SI Imager's integrated gel cutter. Further troubleshooting is discussed in the appendix*

CONCLUDING REMARKS AND FUTURE DIRECTIONS

The SI gel imager is a new fluorescence imaging technology that significantly increases the dynamic range of protein detection in 2DE polyacrylamide gels upwards of 100-fold. By assembling intensity data from multiple exposures, the SI gel imager allows for the detection of a 1,000,000-fold concentration range of fluorescently-labeled protein in a single high-dynamic-range image. The SI gel imager's high dynamic range enabled the detection of the entire resolvable protein concentration range of *Drosophila*. These advances make the SI gel imager a very useful tool for peering more deeply into the proteomes of complex samples, identifying small changes in protein abundance, mass and charge, changes that are often reflective of

physiological processes.

As with all experimental system, there is room for improvement and limitations. The limit of detection (LOD) of the SI gel imager is approximately 10 pg of 10X-labeled protein. One of the factors influencing the LOD is the intrinsic background fluorescence of the system. The background fluorescence signal is about 45 CPS. Most of this signal, 30 CPS, is due to the BBAR anti-reflection –coated, fused silica plates upon which the gels were mounted. We found that the BBAR coating was not stable over long periods of use. A more rugged anti-reflection treatment may lower the background fluorescence. Further engineering of the imager's interior may also reduce the background due to light scatter.

Given the current background signal, the longest possible exposure is about 600 s. At such long exposures the background signal will exceed the masking threshold, thus all pixels will be masked.

The extent of dye labeling plays a major role is influencing the SI gel imager's LOD. The LOD for 1X-labeled ADH was about 1 ng, while 10X-labeled ADH allowed for the detection of as little as 10 pg of ADH. In 1X-labeling approximately 5% of all protein molecules carry a 0single dye molecule, the rest are unlabeled. In 10X-labeling, upwards of 50% of the proteins carry one dye molecule. However, there will be populations of proteins that carry two or more dye molecules. This heterogeneity may negatively impact gel resolution and spot detection. An alternative labeling scheme is to use cysteine-reactive, maleimide dyes–referred to as saturation labeling where all cysteines are labeled. Preliminary experiments with Cy3- and Cy5-maleimide yielded exceptional bright fluorescent protein spots. With an eye toward MS identification of the proteins isolated on 2DE gels, we loaded 100 µg of cellular extract per 2DE gel. The combination of this amount of protein with saturation labeling lead to obvious dye:dye artifacts

where the dye was quenched in regions of high abundance proteins and proteins that had a relatively high cysteine content (Fig 1-18). Finding a labeling scheme where all proteins within an extract carry one and only one dye molecule would be idea for maximizing the LOD of the SI gel imager.

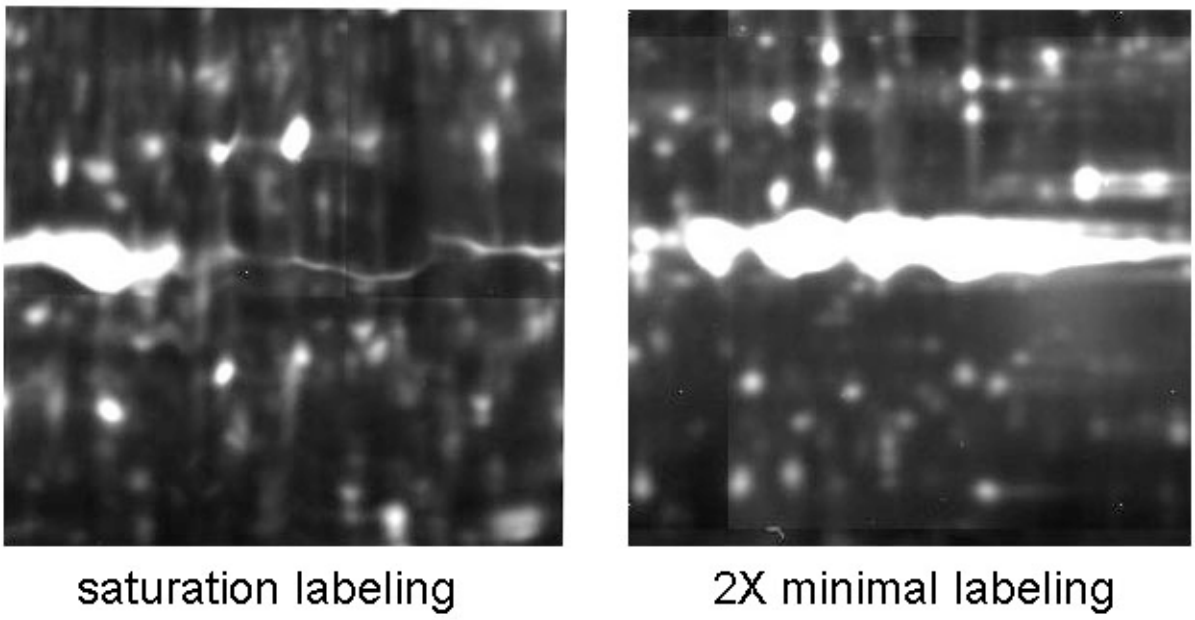


Figure 1-18 *Dye quenching observed in saturation-labeled DEE 2DE gels. The same area of two 2DE gels containing DEE samples is shown.* ***Left****: the saturation-labeled sample experienced dye quenching at high protein abundance area of the yolk proteins.* ***Right****: By comparison, minimal-labeled samples do not experience quenching even up to 10x excess labeling with minimal dyes (2X-labeled sample shown).*

In spite of these limitations, the SI gel imager is still capable of detecting 10 pg of fluorescently tagged protein. For an average sized protein of 50 kD, this amount of protein

represents 0.2 fmol. Current methods for MS identification of proteins isolated from acrylamide gels require about 10 fmol. Now that the SI gel imager is able to detect sub-femtomole quantities of protein, there is a need to develop methods to improve the transfer minute amounts of proteins from an acrylamide gel to a mass spectrometer.

As a developing technology, there remain many areas of improvement for the SI gel imager. Improvements in projector technology may one day result in a projector array with perfect extinction, completely avoiding the "terracing" problem discussed earlier. The current video project used in the SI gel imager is a LCD-based array with a contrast rating of 1:1000. DLP projectors, with contrast ratings of 1:5000, would certainly improve the masking process.

For visualizing high-dynamic-range images, more sophisticated local intensity displaying can make protein spots more distinguishable in CPS images, though our current approach of simply changing the display range in SILab or ImageJ is adequate for most purposes. Spot quantification can be improved by using a combination of commercial software packages or a rewriting of SExtractor for 2DE gel spot detection given its open-source nature. We have initiated collaboration with Kovacivek's group to this end.

Since SILab was written incrementally onto an existing codebase, many instances of inelegant programming and redundancy could be eliminated by a complete rewrite of the program using modern software tools, and certain functions (*eg.* gel cutting) could be refactored and externalized to external script files rather than be hard-coded in the main executable to make SILab more adaptable. SILab's GUI can be refined to reduce operator error.

The fused-silica gel plates currently in use with SI Gel Imager provides very low fluorescence background signal, chemical inertness and mechanical strength sufficient for handling 2DE gels but they are very expensive, ~$700 per plate. We have investigated different

materials for gel plates that has one or more of the following improvements over our fused silica setup: (1) cheaper while still commercially available, (2) lighter/thinner but comparably tough, (3) has lower fluorescence background. No one material so far has had all these characteristics: GorillaGlass is not commercially available for research; a proprietary glass formulation used to cover Apple laptops is thinner, cheaper and with comparable fluorescence background but proved disappointingly fragile. We are still pursuing alternative glass options.

Tangentially, the imager's integrated robotic gel cutting syringe can also be improved. Currently the cutter is prone to incomplete cuts with gradient gels due to the difference in gel hardness. Our current gel-cutting workflow requires cooperation with an external mass-spectrometry core facility, resulting in long wait times between gel running and protein identification in which low-abundance proteins could be degraded or contaminated. A more optimal approach we are investigating is to electroelute the protein spots directly into a trypsin-containing agarose medium for digestion. Proof-of-concept experiments supporting this idea were performed by Lina Song and Brendan Redler during their rotation projects in 2013.

CHAPTER 2: IMPROVING 2DE WITH IN-GEL PROTEIN EQUILIBRATION

INTRODUCTION

Two-dimensional electrophoresis (2DE) is a powerful proteomic technique, capable of resolving large numbers of proteins (Zabel et al., 2009) with very high spatial separation by pH and molecular weight. As a result it has benefited from steady improvements such immobilized pH gradients (Bjellqvist et al., 1993), gradient gels (Walker, 1984) and improved fluorophores and been used to study a wide variety of biological samples. Current 2DE gel workflows involve separating the complex protein sample by performing isoelectricfocusing (IEF) in Immobilized pH Gradient (IPG) strips. The strips are then equilibrated by being immersed in a solution of low concentration urea and high concentration SDS to exchange the urea/CHAPS non-ioinic denaturants for anionic SDS. This is done in two steps: the first includes dithiothreitol (DTT) to reduce sulfhydril groups, and the second step is done in a solution of iodoacetamide (IAA) to permanently alkylate the sulfhydrils, both these equilibration solutions contain SDS. After equilibration, the strips are transferred to large-format polyacrylamide gels where the proteins are separated in the 2nd dimension by SDS-PAGE. The equilibration steps are performed in preparation for subsequent identification using tandem mass spectrometry (MS/MS). Without equilibration, MS/MS is complicated by free sulfhydryl groups' formation of anomalous polypeptide chains. Unfortunately, up to 30% of initial starting proteins have been reported to be lost during equilibration (Zhou et al., 2005), due mostly to diffusion of proteins out of the IPG strips and into the equilibration solutions.

We report here a refinement to the equilibration process, illustrated in Figure 2.1, in which the protein sample is electrophoresed through agarose stacking layers on top of the large format 2nd-dimension SDS-PAGE gel.

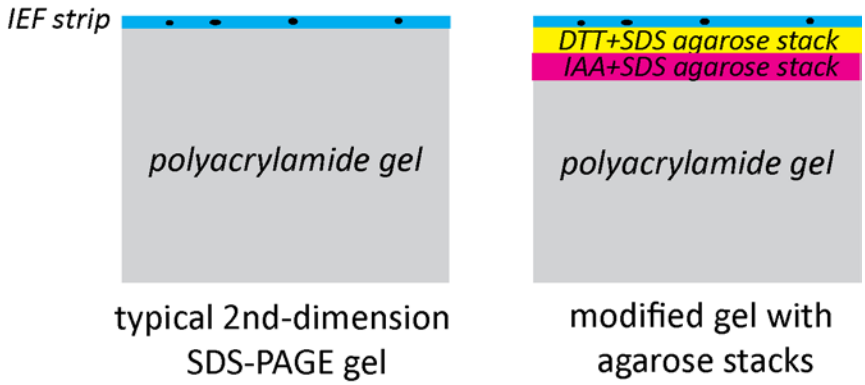


Figure 2-1. The *structure of agarose stacking gels used by in-gel equilibration. For our experiments, 1% low-melting SeaPlague agarose was used. Additional experimental details are found in the text.*

Since the stacking layers contain SDS plus DTT and SDS plus IAA, respectively, the protein samples undergo equilibration without risk of diffusion. This approach is effective at equilibration while reducing sample loss in 2DE gels by minimizing protein diffusion into the equilibration solutions.

MATERIALS AND METHODS

Validation of In-gel Reduction and Alkylation in 1-dimensional, SDS-PAGE Mini-gels

We characterized reduction and alkylation of our in-gel equilibration method using purified Bovine Serum Albumin (BSA, New England Biolabs) since it contains 17 sulfhydryl-containing cysteine residues. As loading controls, BSA was pre-labeled with Cy5-NHS (GE Healthcare, Uppsala, Sweden) as per manufacturer's directions. Two BSA-Cy5 samples in Laemmli sample buffer were prepared with or without β-mercaptoethanol (βME).

Three small-format 12% polyacrylamide gels were prepared: a control SDS-PAGE gel (Figure 1, top panel), a reduction gel, and a reduction-and-alkylation gel. The control SDS-PAGE gel contains only 12% polyacrylamide, the reduction and reduction-and-alkylation gels also contain agarose stacks comprised of 1% low-melting SeaPlague agarose (FMC BioProducts,

Rockland ME), 65mM TRIS and 0.1% SDS. The reduction gel containing a shortened polyacrylamide section topped with a 1 cm high stacking gel composed of agarose containing SDS and DTT (middle panel) and 1 cm high top polyacrylamide section containing the sample loading wells. The reduction and alkylation gel containing a further shortened polyacrylamide topped with a 1 cm high agarose stacking layer that contained 65mM DTT, a 1cm high an agarose layer containing 130mM IAA and a 1 cm high top polyacrylamide section containing the sample loading wells (bottom panel).

1 µg of Cy5-BSA was loaded into each gel lane, comparable to the most abundant proteins in a typical 2DE gel. Electrophoresis was done in standard Tris-Glycine-SDS buffer, running at 60 V through the stacking gels (~15 minutes), and then at 130 V until the dye front had run off the bottom of the gels, another 2 hours. After electrophoresis, the polyacrylamide portion of the gel was then post-electrophoresis labeled with Cy3-maleimide dissolved in Tris buffer for one hour, and then washed in destain (40% methanol, 10% acetic acid) for one hour. The gels were then imaged with matching Cy3 or Cy5 filters and saved as 16-bit TIFF images. For quantification, rectangular selection boxes were manually drawn around the gel lanes using ImageJ and the total integrated flux was measured.

Quantification of Protein Loss Throughout the 2DE Workflow

To quantify protein loss in the transition from IEF to SDS-PAGE, 2DE gels of *Drosophila melanogaster* embryo protein extracts with different equilibration conditions were evaluated. Wild-type Drosophila Embryo Extract (DEE) was obtained as previously described (Gong et al., 2004). DEE protein concentration was determined using a Bradford colorimetric assay (Bio-Rad, Hercules, CA) and diluted with deionized water to final concentration of ~2

μg/μl. For each comparison, two 50 μg samples of DEE were labeled with 1 nmol of Cy5-NHS (GE Healthcare, Uppsala, Sweden) on ice for 30 minutes and quenched with methylamine for 45 minutes. IEF was done using pH 3-10NL, 18 cm Immobilized pH Gradient (IPG) strips (GE Healthcare) on an IPGphor (Amersham Biosciences, UK) for 32,000 Volt-hours. Immediately after finishing IEF, the IPG strips were imaged using Cy5 filters for 1 second exposures to prevent detector saturation, since the plastic backing of the IPG strips had very low background fluorescence signal in the Cy5 emission range. Quantification of total resolved protein was done by drawing a selection box encompassing the central three-fourths of each strip with ImageJ and the integrated flux was measured. The regions outside of the central portion of the IEF strips contained protein that precipitated at the loading well and where the bulk of the free dye accumulated.

Second-dimension SDS-PAGE was done on 20x20 cm, 12% polyacrylamide gels at 4 °C in standard Tris-Glycine-SDS buffer. The solution-based equilibration IPG strip was rinsed in a solution containing 2% w/v SDS, 6M urea and 65mM DTT for 15 minutes, rinsed 3 times in deionized water, imaged and transferred to polyacrylamide gel. For the in-gel-equilibration of the IPG strip, an agarose stacking gel containing 2% SDS and 65mM DTT was poured on top of the polyacrylamide gel and allowed to completely harden, approximately 15 minutes, before the strip was added on top. It is important to ensure that the top of the polyacrylamide gel is completely dry before pouring agarose to ensure even equilibration across the strip. Both strips were sealed with 1% SeaPlague agarose containing 2X Bromophenol Blue tracking dye. The second dimension gel was electrophoresed at 60V until the tracking dye front is past the agarose stack (approximately 1 hour), then at 300V until the tracking dye front was completely off the gels, approximately an additional 5 hours. We find that exceeding 300V generates poor quality gels,

likely due to excess heat. Using these running times, we determined the optimal height of the agarose stacks to be approximately 1.2 cm. After 2nd dimension separation, both IEF strips were then removed from the gels, imaged and quantified as described above.

After gels have been in destain for one hour, Cy5 images of the gels were obtained using the Structured Illumination Gel Imager (SI Gel Imager) fluorescence imager and saved as 16-bit TIFF images. To quantify spots, the central area of each TIFF was submitted to DeCyder Difference Image Analysis (DIA,(Kang et al., 2009a)) version 5.0, set to detect 400 protein spots in an unsupervised fashion. DIA detection results were visually inspected and spurious detections (*eg.* dust particles, irregular shapes unlikely to be genuine protein spots) were manually excluded from further analysis. To compare DeCyder-reported spot volumes (defined as the sum of all pixel intensities within the boundary of a spot), 21 verified spots common across in-gel-based and solution-based equilibration experiments were manually selected. Pearson's correlation for spot volumes was obtained in Microsoft Excel (Microsoft, Redmond, WA). A total of 5 in-gel equilibration gels were compared to 5 in-solution equilibration gels.

RESULTS & DISCUSSION

Validation of In-gel Reduction and Alkylation

We first wanted to establish the feasibility of in-gel equilibration using a known quantity of protein with a high abundance of –SH groups. To determine if stacking gels composed of agarose containing DTT and IAA were sufficient to first reduce and then alkylate proteins as they migrate through the stacking media, a series of SDS-PAGE gels were run with different stacking gel compositions. Agarose was chosen as the desired gel media because DTT and IAA

inhibit acrylamide polymerization. The results of these in-gel reduction and alkylation experiments are shown in Figure 2-2.

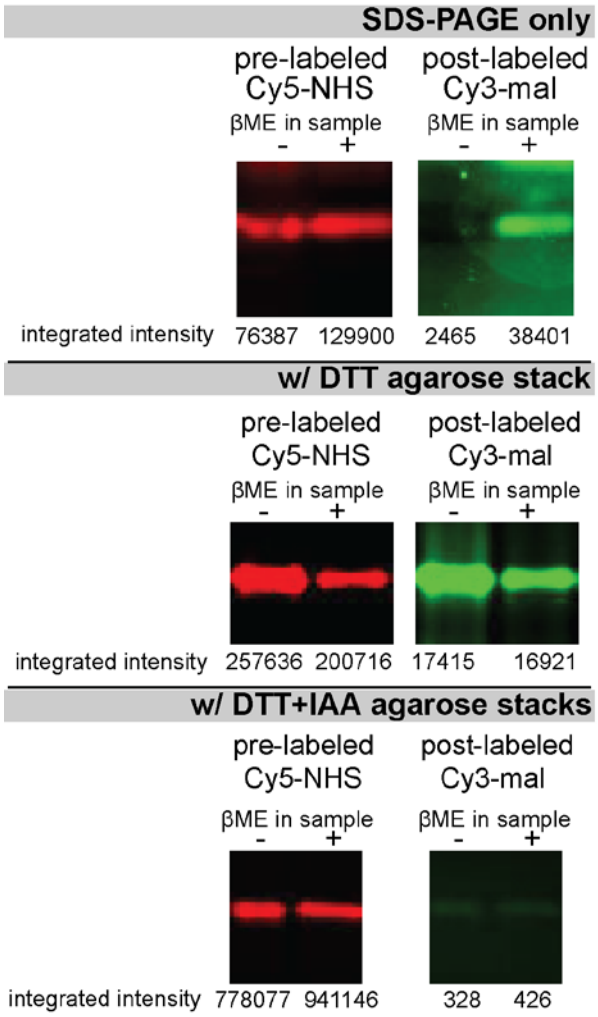


Figure 2-2. *In-gel equilibration of purified protein standards in 1D SDS-PAGE mini gels. BSA was pre-labeled with Cy5-NHS as a loading control and electrophoresed through 12% polyacrylamide gel (top row), polyacrylamide gel with DTT-containing stacking agarose gel (middle row), or DTT-containing and IAA-containing stacking agarose gels (bottom row). All gels were post-labeled with cysteine-reactive Cy3-Maleimide to probe for the presence of reduced –SH groups, imaged and quantified.* ***Top:*** *Cy3-Maleimide labeling was posible only when samples were initially reduced with βME.* ***Middle:*** *An agarose stacking gel containing DTT*

reduced –SH groups enabled maleimide labeling even in the absence of βME in the sample buffer. ***Bottom:*** *Alkylation of –SH groups by IAA in the stacking gel prevented post-labeling even after reduction by DTT in the first agarose stacking layer.*

When reducing βME was present in the sample, the reduced –SH groups of BSA were able to react with cysteine-reactive Cy3-Mal (top panel). In the absence of βME in the protein sample buffer, the BSA was not labeled with Cy3-Mal. When DDT was introduced into the agarose stacking gel media, both βME-containing and βME non-containing samples became reduced and were labeled with Cy3-Mal (middle panel). The extent of reduction appeared to be equivalent between the samples prepared with and without βME since the relative fluorescence intensity between the Cy5-NHS pre-labeled BSA and the Cy3-Mal post-electrophoresis labeling was relatively the same. The addition of an IAA-containing, alkylating stacking gel layer between the DTT stacking gel and the polyacrylamide resolving gel demonstrated that the reduced proteins could be extensively alkylated as evidenced by the near complete inhibition of Cy3-Mal labeling of the BSA after electrophoresis (bottom panel). Taken together, these results show that in-gel reduction and alkylation can be readily accomplished by incorporating DTT and IAA into agarose stacking layers. In addition, the agarose stacking layers appear to function as effective stacking media since the BSA protein bands remained well-defined bands.

Quantifying Protein Loss Throughout the 2D Electrophoresis Workflow

To test our hypothesis that in-gel equilibration allows for better sample retention, we compared the amount of sample that remained detectable within the IEF strip at each step of the 2DE workflow, comparing solutions-based and in-gel equilibration schemes. The fluorescent signal detected after major steps of the workflow: after IEF, after in-solution equilibration, and

after SDS-PAGE (Figure 2-3).

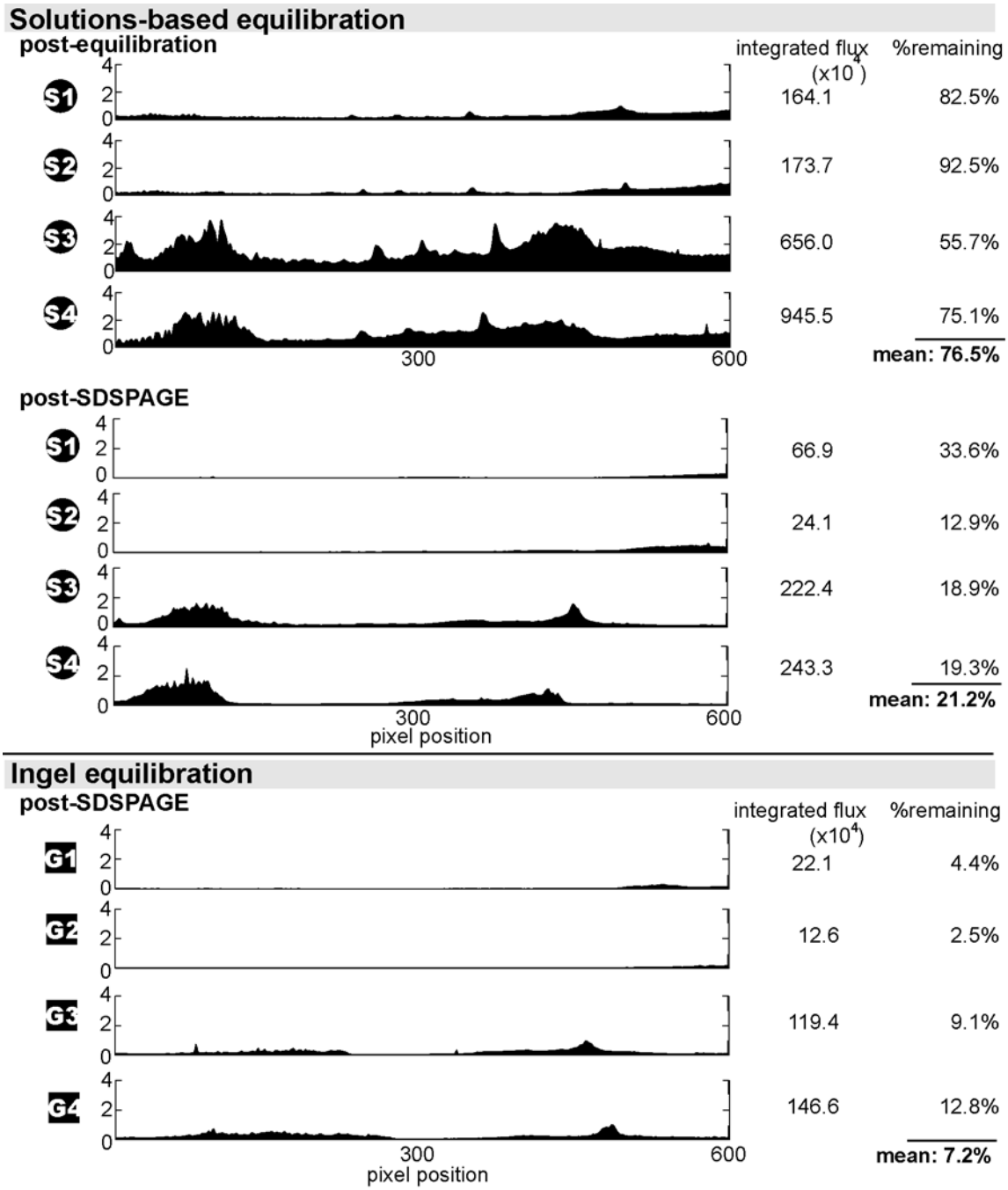


Figure 2-3. *In-gel equilibration increased sample retention between IEF and SDS-PAGE. Four solutions-equilibration (S1 – S4) and four gel-equilibration samples (G1-G4) of DEE labeled with Cy5-NHS*

Cy5-labeled protein was used for these experiments to avoid the shorter wavelength auto-fluorescence of the IEF strips. Four solution-equilibrated samples (S1 through S4) were quantified post-equilibration and post-SDS-PAGE. On average, about 77% of the initial sample remained after equilibration, consistent with previous studies that show up to 30% sample loss due to in-solution equilibration (Zhou et al., 2005). In protocol described here used only a single 15 minute equilibration step in an equilibration buffer containing DTT, the second equilibration step with IAA was omitted, which may further reduce the amount of protein remaining in the IEF strip after equilibration. After SDS-PAGE, approximately 21% of the initial sample still remained in the IEF strip, representing further sample loss due to protein remaining within the IEF strip matrix after second dimension electrophoresis.

In-gel equilibration eliminates the in-solution equilibration step. Therefore, there was no opportunity for proteins to diffuse out of the gel during equilibration. In important concern was that the exchange of first dimension denaturants, urea and CHAPS, for the second dimension denaturant SDS during electrophoresis out of the IEF strip, through the stacking gel, and into the SDS –PAGE resolving gel would be sufficient. Imaging of in-gel-equilibrated IEF strips (G1 through G4) after second-dimension electrophoresis showed that an average of 7.2% of the initial starting samples remains within the IEF strips, demonstrating that more than 90% of the resolvable protein left the IEF strip and entered the second dimension gel. This is in contrast to in-solution equilibration where about 55% of the initial sample eventually entered the second dimension gel (24% of the initial protein sample was lost during in-solution equilibration and 21% of the initial sample remained within the EIF strip after second dimension electrophoresis). Thus in terms of bulk protein, the in-gel equilibration protocol was significantly more efficient in retaining the sample from IEF to SDS-PAGE.

Comparison of Post-equilibration 2DE gels of Complex Samples

To test if this almost two-fold loss of bulk protein seen in the in-solution equilibration compared to in-gel equilibration translated into similar losses for individual proteins, we compared protein spot abundance changes in the large-format 2DE gels. We compared a subset of 20 spots between five in-solution- and in-gel-equilibrated gel pairs. To minimize sample-to-sample variation and IEF-to-IEF strip variation, the same DEE sample was used for each pair of IEF strips, where one strip was in-solution equilibrated and the other strip was in-gel equilibrated. Each pair of strips was from the same batch of strips. For each protein spot, spot volumes were calculated by DeCyder and the in-gel/in-solution spot ratio determined. Comparisons between two gel pairs are shown in Figure 2-4. For the majority of the spots, the in-gel counterpart had higher spot volume (gel/solution spot volume >1.0), and the spots were more visible in the image (panel A). This improvement was seen over 5 different replicate gel pairs (median ratio > 1.0, panel B). The average increase in spot volume (intensity) was 1.8-fold. This increase in spot volume closely corresponds to the measured 45% loss of bulk protein seen for the in-solution equilibrated strips.

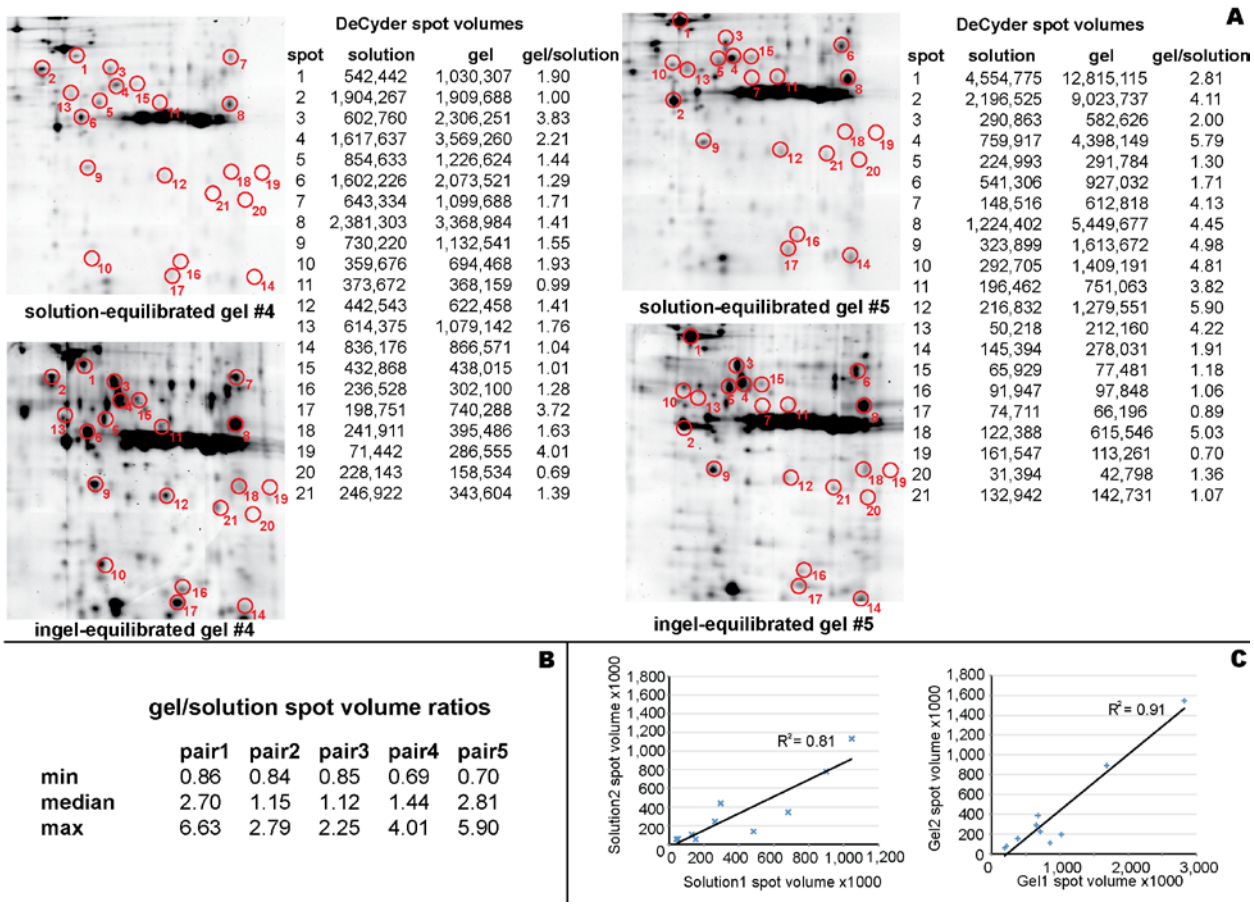


Figure 2-4. *In-gel equilibration improves spot retention and reduces gel-to-gel variation for 2DE experiments. (A) Two representative pairs of 2DE gels (pair4 and pair5) comparing solution- and in-gel-equilibrated samples of DEE are shown. Numbered circles are spots common to each gel pair being compared; spot volumes are reported by DeCyder. (B) Summary of spot volume ratios across five replicate gel pairs. (C) Spot volumes of the same ten spots are correlated between two solution-equilibrated samples (left panel) and between two in-gel-equilibrated samples (right panel).*

While the average protein increased volume when comparing in-gel to in-solution equilibration, there was a wide range of variability between these pair-wise comparisons. We hypothesized that in-gel equilibration would enable proteins in the sample to migrate into the 2nd

dimension SDS-PAGE gel with better uniformity than in-solution equilibration, thus reducing some of the gel-to-gel protein spot variability seen in when running 2DE gels using solution equilibration. To compare gel-to-gel spot variability, 10 common protein spots were selected from two solution-equilibrated gels and their volumes plotted against one another, giving a Pearson's R^2 value of 0.81 (Figure 2-4C, left graph). In contrast, the same 10 spots across two in-gel-equilibrated gels had an R^2 value of 0.91, indicating lower gel-to-gel variability for in-gel equilibrated gels (Figure 2-4C, right graph).

In conclusion, these experiments have demonstrated that directly transferring an IEF strip immediately after the completion of isoelectric focusing to a second dimension SDS-PAGE gel which has agarose stacking layers that contain DTT and IAA provides for a more efficient and consistent transfer of protein across the two separation axes of 2DE gels. In our experience, the transition between the first and the second dimension of 2DE poses the largest technical challenge for novice 2DE practitioners. The development of this in-gel equilibration procedure greatly reduces the potential errors associated with in-solution equilibration. Moreover, the elimination of in-solution equilibration provides a more facile route for automating the running of large-format 2DE gels.

CHAPTER 3: MANAGEMENT OF UNDERGRADUATE RESEARCH TEAMS

The Need for a Proteomics "Platoon"

Increasingly undergraduate students are expected to have realistic laboratory experience before continuing onto later stages of their careers. While the relevance of working in a university research laboratory is apparent for students bound for the biomedical fields, skills in communication, negotiation, planning and ultimately leadership are broadly beneficial for all undergraduates regardless of future involvement in research.

Challenges involved in fostering undergraduate research include: scheduling, maintaining student interest, productivity, and presenting or publishing their research. To address these challenges, the Minden Lab devised a system that we call the "Proteomics Platoon" in which undergraduate students mentor each other in independent two- or three-person teams to conduct proteomic experiments.

Proteomics experiments require several days to complete and no single undergrad has time in their busy schedule to attend to all of the steps of an experiment to complete it effectively. The platoon model allows the various members of the platoon to assist each other and train each other to complete their tasks as a coordinated unit.

Goals of the Proteomics Platoon

The major objectives for the members of the Proteomics Platoon are: (1) to gain technical competency across all experimental steps, (2) learn to communicate and cooperate in a professional scientific workplace and (3) learn current biology by working with realistic samples and experiments. Since our laboratory has close collaborations with other research groups, our

goals also include making progress reports to our collaborators, striking a harmonious balance between undergraduate education and scientific productivity.

Platoon Recruitment and Comparison with CMU Laboratory Classes

To date our students have come from CMU Biology or Chemistry and were introduced to the laboratory through word of mouth or CMU Colloquia or Seminars attended by CMU Biology sophomores. Near the end of the Fall Semester (after students have requested to join the platoon and are vetted by Dr. Minden and the rest of the platoon), our laboratory holds a platoon meeting to introduce new platoon members and discuss projects. Students join the Platoon at the beginning of their sophomore year and typically remain involved through graduation. Beginning in their second year ensures that students have had exposure to chemical and biological safety training and are familiar with the workings of the Department, while providing enough time for meaningful student development, typically resulting in peer-reviewed publications.

Despite their similar goals of endowing undergraduates with research experiences, the Platoon system and CMU Biological Sciences laboratory classes (required for all CMU Biological Sciences majors, and optional for CMU Chemistry majors) are different in several important ways. Four platoon members were interviewed in March 2013 in which they were asked to describe these differences, summarized in Table 3-1

CMU Undergraduate Laboratory Courses	**The Proteomics Platoon**
- Protocols are established and static for the entire semester	- Protocols are rapidly adapted based on experimenter feedback
- Experiments are scheduled in advance for the entire semester	- Experiments are scheduled as needed at the beginning of each week

- Scientific background is established, and provided in the student's curriculum in advance of the laboratory class	- Scientific background is often in primary literature, lab records, or scarce (for nascent experiments)
- Experiments are conducted in 3-hour preset lab periods	- Experiments often last days or weeks, performed at irregular interval depending on availability and sample preparation
- Each experiment type is typically conducted once in the semester (*eg*. gel electrophoresis)	- Experiments are repeated under successful or sufficiently replicated
- Students are equal partners	- Incoming members are mentored by more experienced members
- Results are presented in written lab reports with prescribed format	- Results are orally presented at lab meetings, scientific meetings and publications
- Student performance is gauged by regularly-assigned assignments, tests and quizzes	- Performance is gauged by experiment success and member competence

Table 3-1. *The Proteomics Platoon compared to typical CMU laboratory courses*

Types of Experiments Conducted by the Proteomics Platoon

As detailed in Chapter I, the construction of SI gel imager enabled high dynamic range gel imaging in the Minden Lab. As a result, SI gel imager benefited from having real-world samples in the form of 2DE gels for testing the imager in its early development stages. The bulk of experiments therefore involve running 2DE gels, which can be further broken down into the following tasks:

- Measuring the concentration of protein samples using the Bradford Assay

- Labeling proteins with fluorescent NHS-ester dyes. The vast majority of the Platoon's experiments utilize Difference Gel Electrophoresis (DIGE, (Unlu et al., 1997))

- Rehydrating Immobilized pH Gradient (IPG) strips in preparation for isoelectric focusing (IEF)

- Loading labeled proteins into the IEF apparatus and starting IEF

- Pouring large-format, second-dimension, SDS-PAGE gels

- Transferring IPG strips onto SDS-PAGE gels

- Second-dimension Electrophoresis

- Obtaining fluorescence images with SI Gel Imager

- Analysis of 2D DIGE results

Timing for these tasks is outlined in Figure 1, tasks of uncertain or variable time requirements indicated by gradient shading. More senior platoon members also take part in project-specific experiments as needed and permitted by time and availability of mentors in the laboratory:

- Spot cutting: excising a polyacrylamide gel fragment using an in-house robotic cutter

- Cell culture: counting, passing and preparation of cell lysates

- Fruit fly genetics: sexing flies, identifying phenotypes with light microscopy, preparation of protein extracts from fly embryos

- Molecular cloning and western blotting

Group Formation and Assignment of Projects

The platoon is subdivided into groups by projects, each group consisting of a senior student (typically CMU junior or senior) and a junior student (CMU sophomore). Overall project directions are set by the faculty advisor in conjunction with outside collaborators.
Daily experimental planning and conduct is overseen by the lab's graduate students.

Project & collaborator	**Platoon members**	**Duration**
1 Acute myeloid leukemia(AML)	Victor Bass*, *Alex Rodriguez*	06/2012 –

Dr. Mark Minden, University of Toronto		06/2013
2 Rheumatoid arthritis-interstitial lung disease (RA-ILD) Dr. Dana Ascherman, University of Miami	*Minh Le*, Raghunandan Avula*, *Dagney Cooke**, Alex Hurley (NSF REU)	08/2012 -
3 The role of APC1 and 2 in Drosophila embryo development (APC) Dr. Brooke McCartney, CMU	*Amritha Parthasarathy**, Danielle Schlesinger*, Lawton Tellin*, Taylor Maggiacomo*, Anna Pyzel*	06/2012 -
4 Plant root proteomics (*Arabidopsis*) Dr. Philip Benfey, Duke University	Victor Bass*, Rachel Willen*	07/2012 -
5 Mitochondrial changes associated with Huntington's disease (Huntington's) Dr. Robert Friedlander, University of Pittsburgh	TBD	05/2013 -
6 Starfish embryo regeneration after bisection (Starfish) Dr. Veronica Hinman, CMU	TBD	05/2012 -

Table 3-2. *Projects undertaken by the Proteomics Platoon, 2011 – 2014. Detailed project descriptions are in the text. Italics indicate students are no longer in the platoon as of 03/2013. (*) indicates undergraduates associated with the HHMI Summer Research program.*

Project 1 (AML): Acute Myeloid Leukemia is a form of blood cancer normally treated by taking the enzyme Asparaginase. The platoon members are running DIGE gels comparing proteomes of AML cultured human cells at different time points in response to Asparaginase treatment.

Project 2 (RA-ILD): Rheumatoid Arthritis with Interstitial Lung Disease is a fatal condition with no known cure. A hallmark of RA is the increased level of arginine modification of cellular proteins into citrulline, which leads to the production of anti-citrulline auto-antibodies. This project intends to identify auto-antigen differences between RA, RA-ILD and healthy patients' cells. As the samples have been collected through antibody affinity assays, the platoon will be working in conjunction with Vinitha Ganesan to purify the samples using a reversible protein clean-up method currently being developed in the lab. The platoon will be running DIGE gels comparing the purified samples.

Project 3 (APC): Adenomatous Polyposis Coli is a protein that regulates Wnt signaling. Since Wnt is a major signaling pathway, disruption in APC function is implicated in many pathologies, including colon cancer in humans. The McCartney Lab is investigating the effects of APC malfunctions in Wnt signaling using *Drosophila* embryos as a model due to the high homology between *Drosophila* and human APC proteins. The platoon will be running DIGE gels of different APC mutants to identify proteins involved in APC malfunction.

Project 4 (*Arabidopsis*): Due to the relatively slow progress of plant proteomics, we are working on generating high-quality 2DE gels of *Arabidopsis* proteins from shoots and roots to identify proteins involved in differentiation and development of these tissues. Previously the protein extracts we received suffered from low protein concentrations and protease degradation. We will be concentrating the samples and using protease inhibitors, testing the samples on 1D gels before running on 2DE gels.

Project 5 (Huntington's disease): Huntington's Disease is a fatal disease in humans linked to mutations in the Huntingtin gene. There is some evidence that protein importation into mitochondria is impaired in HD patients. The platoon members would be running DIGE gels of proteins extracted from primary and cultured mice cells at different stages of Huntington's manifestation.

Project 6 (Starfish): Starfish is known to be able to regenerate up to half of their body upon dissection or traumatic injury. This project seeks to understand the proteomic changes involved in regeneration. The platoon will be running DIGE gels comparing starfish protein extracts collected at different stages of regeneration.

Scheduling Proteomics Platoon experiments

One challenge of experiments involving 2DE gels is the long and variable time required to get from sample to results. As outlined in Figure 1A, a typical 2DE experiment involving 2 replicate gels typically take up to three days, or longer if the protein sample has to be concentrated, desalted or otherwise processed. This is impossible for any single undergraduate student to finish during the regular academic year while classes are in session. However, the workflow can be broken down into smaller tasks that can be done by one or two students.

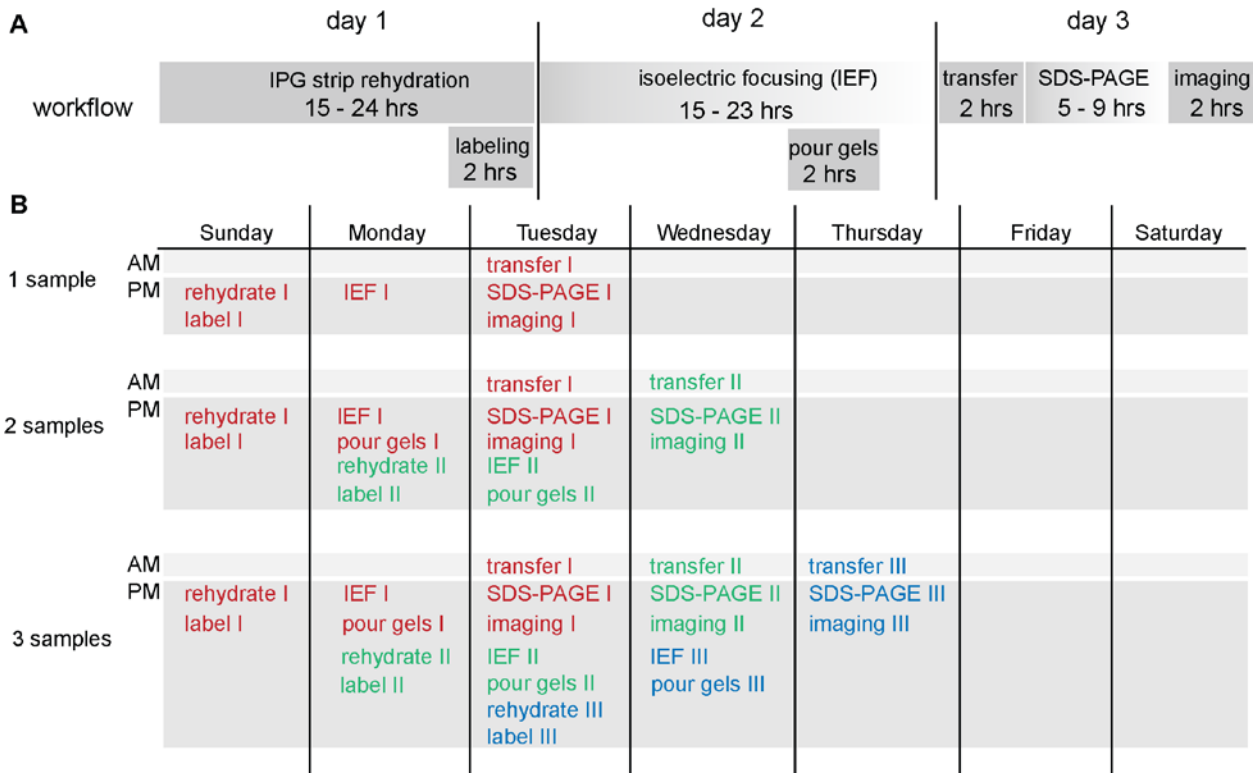


Figure 3-1. *Schedule of experiments for the Minden Lab Proteomics Platoon. (A) The entire 2DE workflow with tasks drawn approximately to scale, gradient shading indicates tasks of uncertain variable duration. (B) A typical week's schedule, allowing for three 2DE experiments*

Tools for Communication Within the Platoon

While the platoon's projects are largely independent, the use of common space, procedures and materials require constant clear and efficient communication between group members. The main tasks requiring communication within the Proteomics Platoon are: scheduling experiments, requesting and fulfilling coverage of experiments and, sharing background literature and experimental results. After experimenting with email, we found that the communication tool most useful to the operation of the Platoon has been GroupMe (Microsoft, Redmond, WA), a mobile phone application enabling single-sender-multiple-recipient SMS text messaging

regardless of phone model or carrier, which enabled real-time announcements of ongoing experiments and rapid request of experiment coverage. For scheduling experiments we found a common experiment calendar much like the one illustrated in Figure 3-1B hosted on Google Calendar provides instant updates and easy access. For record keeping we maintain common traditional paper lab notebooks in compliance with CMU policy. Common protocols, presentations and image data are shared electronically via Dropbox, a persistent online data storage service, with backups stored on the SI Gel Imager computer.

Typical Development of a Platoon Member

The course of a platoon member's development is outlined in Table 3-3. This information was compiled from first-hard observations by graduate students and informal question-and-answer sessions with the platoon members.

Year	**Expectations**	**Typical Challenges**	**Solutions**
Sophomore	Gain fluency with 2DE	Boredom with the training phase Apprehension about destroying precious samples	Balance watching experiments with doing experiments to build confidence and interest
Junior	- Mentor sophomore students on 2DE techniques - Read primary literature and give journal club talks - Design small experiments	Delegation of responsibilities Providing occasional criticism of platoon mate Maintaining records and schedules	Consulting with graduate students and/or faculty adviser
Senior	- Learn new techniques	Lack of time	Design and complete

	- Design larger experiments		small, self-contained projects

Table 3-3. *Development of a typical undergraduate student in the Proteomics Platoon. Each year is discussed in further detail in the text.*

Sophomore year: New platoon members shadow 2-3 2DE experiments over a period of 3-6 weeks, followed by 1-2 weeks of running experiments using test samples. Shadowing is repetitive, and platoon members report preferring doing experiments themselves after the 3rd experiment. Students are expected to identify the problems that occur during the first 2DE experiments that they run themselves, during informal guided "post-mortem/trouble-shooting" conversations at subsequent lab meetings. These conversations emphasized the importance of record-keeping, creativity and sound scientific reasoning without pressures associated with graded exams, and most students reported having found these exercises helpful.

During the summer term between sophomore and junior years many platoon members remain in the laboratory as Summer Researchers or Summer Scholars of the CMU Howard Hughes Medical Institute (HHMI) Summer Research program. This provides an opportunity for the students to perform the entire 2DE workflow without substantial supervision, and start developing presentation skills in preparation for their HHMI presentation. –they gain ownership of their projects over the research-intensive summer projects.

Junior year: Platoon members attend new project meetings and begin mentoring new platoon members. Students typically become frustrated with lack of progress, having to balance mentoring newly-trained sophomore students while working on their project. Students are encouraged to gather information about their projects from primary literature (or research grant

documents for new projects) then explain them to new platoon members. This process stresses the importance of sound scientific understanding and further hones communication skills learned in the previous summer.

Senior year: Platoon members are autonomous, capable of designing experiments independently and needing only to coordinate schedules with the rest of the platoon. Student involvement tapers off in the second semester due to post-graduation preparations, typically interviews for graduate or medical school.

Fostering Leadership in Junior Platoon Members

In the "best practice" of [undergraduate research], the student draws on the "mentor's expertise and resources. . . and the student is encouraged to take primary responsibility for the project and to provide substantial input into its direction" (Hunter et al., 2007). It has been commonly observed that in large groups of undergraduate researchers, a small number rise to assume leadership roles (Whiteside et al., 2007) .The Platoon system develops this idea further by explicitly requiring that senior members mentor incoming junior members. Most, though not all students develop along this path, few have left the platoon due to a lack of sustaining interest or increased external time demands.

Benefits for Graduate Students

The Platoon gives graduate students an opportunity to become mentors, coordinators and mediators. We believe the platoon system confers to graduate students three important skills. First is the ability to provide leadership by example. Since platoon members begin their involvement with minimal laboratory experience, the platoon system requires the graduate

students to be competent in demonstrating the techniques to platoon members while also possessing a degree of patience in trusting platoon members to eventually become self-sufficient. Leadership by example also extends to day-to-day attitude, where the graduate students remain calm, friendly and professional in the laboratory regardless of stresses that may arise. The second skill is the ability to balance one's own thesis work and managing the platoon. This requires the graduate students to accurately gauge the platoon's competence and confidence while delegating tasks. The third skill is the ability to motivate undergraduate students, whose time and attention are constantly being divided amongst classes, extracurricular and personal commitments. This requires the graduate students to have reasonable expectations of the platoon's availability while tailoring motivational approaches to individual members of the platoon.

Shortcomings of the Proteomics Platoon and Proposed Solutions

While we feel the platoon system have made progress addressing some problems involved in providing undergraduates with authentic research experiences, there remain shortcomings that could be improved. In particular, the platoon requires a significant initial time investment in which junior platoon members are trained to become fluent in basic laboratory techniques. While a training phase is common to many laboratory employing undergraduate workers, the platoon requires more time to become productive overall. This can be addressed by negotiating with external collaborators to reach a common understanding that the projects involving the platoon proceed at a relatively slower pace. The senior platoon members are largely unable to perform experiments while mentoring junior members. As a result, the platoon also requires a significant ongoing time investment for the supervising graduate students.

APPENDIX

1. Creating a Uniform Agarose Gel for Image Registration Calibration

Materials

Lysis buffer (5 ml):

8M urea	2.4 g	(chemical shelf)
4% CHAPS	0.2 g	(4C fridge door, top shelf)
10 mM Hepes pH 8.0	50 ul of 1M stock	(plastic bottle on bench)
diH2O to	5 ml	(glass bottle on bench)

Labeling mix (7x -> 700uL final volume, enough for Cy3 and Cy5 Eppendorf tubes)
70 ul serum (goat, fly room fridge, top shelf)
630 ul lysis buffer

TCEP (10 mM):
MWt 286.65gl/mol (4C fridge door, middle shelf)
Make 2.87 mg/ml solution: weigh ~1-2mg of TCEP and add H2O to 2.87 mg/ml

Cy3-mal and Cy5-mal stock solutions (10 mM):
1 tube Cy3-mal (-80C box marked "Cy3 Cy5 Saturation")
1 tube Cy5-mal (" ")
Add 30 ul dry DMF/tube (aliquot with syringe from air-tight bottle under main lab hood)

Procedure

1. Thaw out CHAPS (should be at room temp. when making buffer)
2. Make lysis buffer in 15ml conical
3. Make labeling mix in 2 Eppendorf tubes
4. Add TCEP and incubate at 37C for 60 minutes
5. Thaw out Cy3/Cy5-mal tubes straight from -80/dry ice by rubbing them
6. Add 30 ul of DMF into each dye tube
7. Add labeling mix+TCEP and incubate at 37C for 30 minutes

Reaction #	**1**	**2**
Labeling mix	330	330
TCEP	15	15
60' @ 37C		
Cy3-mal	30 (whole tube)	
Cy5-mal		30 (whole tube)
30' @ 37C		

Use 25ul per mini gel

2. Creating a Uniform Polyacrylamide Gel for Post-Capture Image Correction

1. Make mix for 1ug total BSA in 12% acrylamide gel (all amounts below in ml)

	12% gel (10 ml)
diH_2O	3.33
30% Acrylamide	4.0
1.5M Tris 8.8	2.5
20% SDS	0.05
BSA (10 mg/ml)	0.001

2. Polymerize gel for 30 minutes

10% APS	0.1
TEMED	0.004

3. Soak gel in 40mL of 1mM TCEP (114.66mg of TCEP). Shake at 37C in the dark for 1 hour.

4. Make 10mM solution of Cy3-maleimide by dissolving 1 vial of dye (300 nmol) in 30uL fresh DMF to make 10mM solution of Cy3-mal.

5. Add 4uL of 10mM Cy3-mal to petri dish. Shake at 37C in the dark for 1 hour.

6. Destain gel for 15 minutes and image

3. Operating and troubleshooting the Structured Illumination Gel Imager

1. SI Gel Imager components

1a. SILab software

2. SI Gel Imager start-up sequence

3. Imaging

3a. DIGE mode

3b. SI Gel Imager mode

3b1. Performing image registration

3b2. Obtaining CPS image

3b3. Working with CPS images

3d. Troubleshooting

4. SI Gel Imager shutdown sequence

5. Image analysis

6. Gel cutting

5a. Setting up gel plate and cutting head

5b. Imaging

5c. Performing the actual cut

5d. Cleanup

5e.Troubleshooting**.**

7. Uncommon tasks

7a. Resetting gel stage after a crash

7b. Resetting camera after a crash

7c. Lubricating stage lead screws

7d. Changing projector air filter

7e. Changing projector lamps

7f. Cleaning the light dump

7g. Realigning imager optics

1. SI Gel Imager components

The SI Gel Imager has the following hardware components:

- Projector (Hitachi CP-X505)

- Camera (Princeton Instruments 16-bit CCD)

- Gel stage

- Gel cutter

- Filter wheels

- Computer (Dell Precision T1500): has a dual-head video card with DVI outputs, one output goes to projector, one goes to computer monitor

- USB-to-serial hub

2. SI Gel Imager start-up sequence

1. Turn all switches on powerstrip to ON by turning the master switch to ON. There will be a low humming noise.

2. Press POWER on the Hitachi projector. Its power light next to the POWER button will turn from amber to green.

3. Log into computer (SI Gel Imageruser / mi277) and start the SILab program by double-clicking its icon on the desktop.

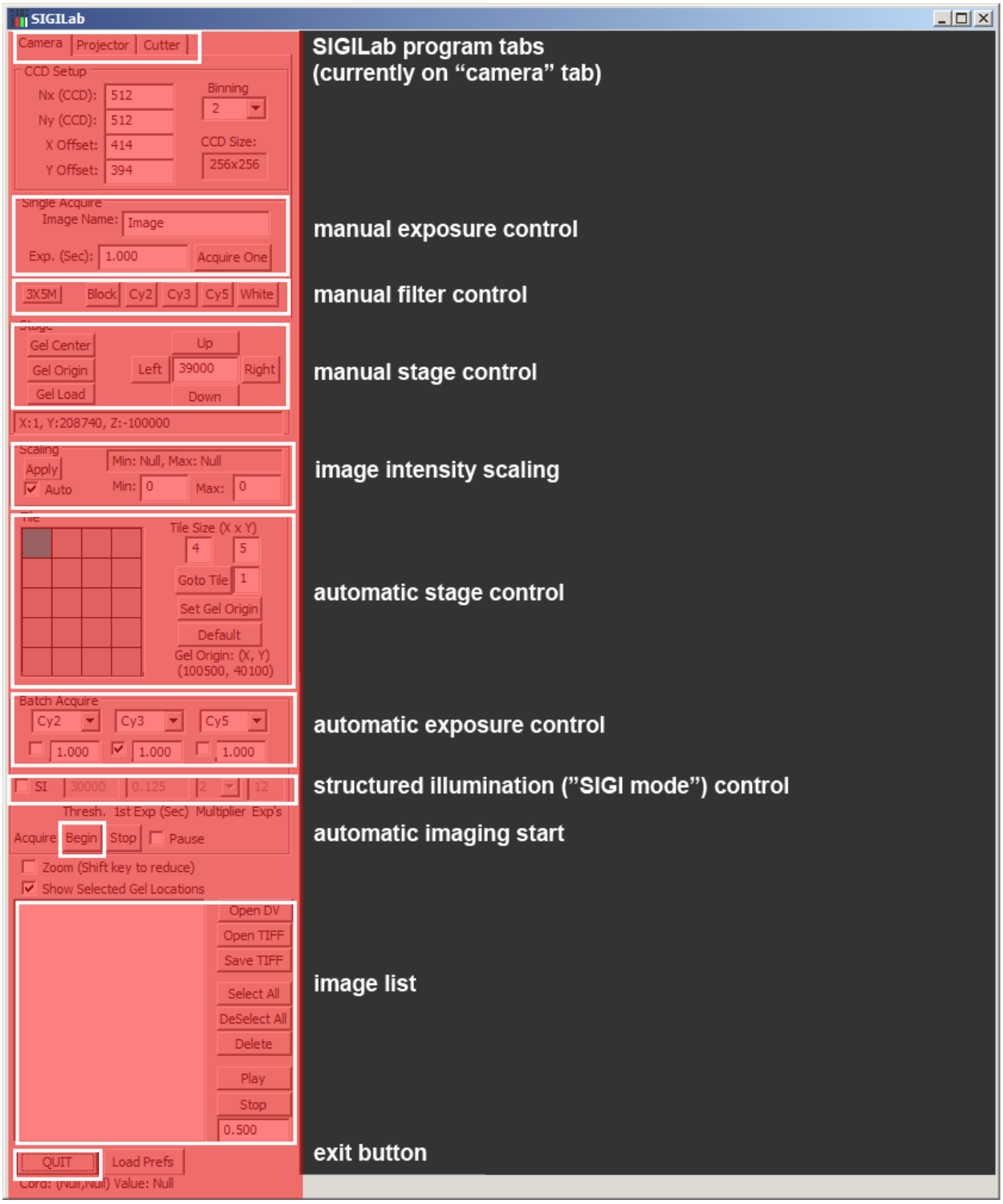
4. SILab will ask whether you want “demo mode” (no hardware will be activated). Choose “No”

5. SILab will ask you to confirm that all hardware (*ie.* gel stage, camera) is turned on. Choose “Yes”. At this point the camera will turn on and start to hum a bit louder.

6. Prepare your sample/gel for imaging. You have 3-4 minutes while the camera warms up

3. Imaging

SILab has a control area on the left where buttons controlling the imager are located, and an image area on the right where images are displayed



The common controls are:

Manual exposure control: taking a single exposure of a single tile

Manual filter control: changes the filters currently in use. Excitation and emission filters are matched and move together automatically

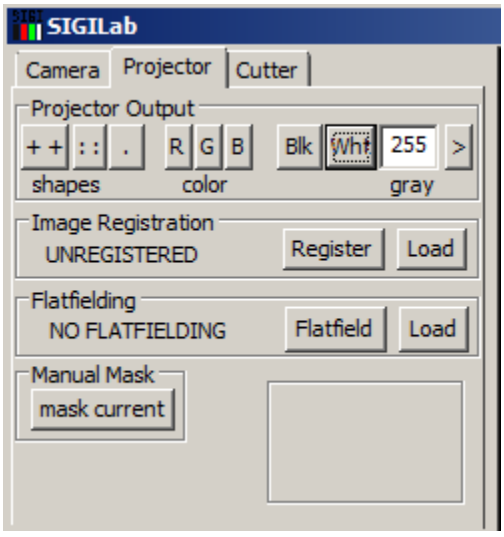
Manual stage control: moves the stage by the distance in the number box (the default 39000 is the size of one tile). Also there are preset buttons for moving to gel center, Gel Origin and gel load positions.

Image intensity scaling: rescale the image between the Min and Max values. Changing Max to values > 0 increases perceived image brightness without altering image data. Also displays the min and max intensity values of the current image

Automatic stage control: changes the size and shape of the imaging area for batch mode (the size of each tile remains the same). The uppermost left tile is Gel Origin, which can be set by clicking “Set Gel Origin”. A typical large-format 2DE gel requires 4x4 tiles. These Automatic stage control settings do not influence the manual stage control settings above.

Automatic exposure control: sets the channels and exposure times for batch mode imaging. SILab saves each channel to a separate file for DIGE mode and all channels to the same file for SI Gel Imager mode. By default Cy3 is selected

Structured illumination (“SI Gel Imager mode”) control: sets parameters for structured illumination, which takes additional exposures for each selected channel selected in Automatic exposure control. This mode requires image registration and SILab will remind user to perform registration if it hasn’t been done (more details below)

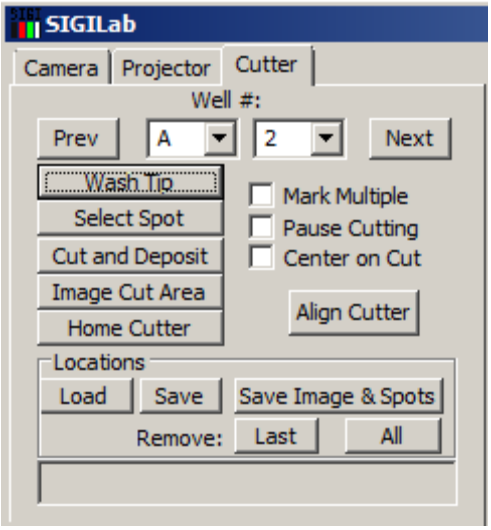


Automatic imaging start: click this button to actually begin imaging using the parameters set in Automatic stage control, Automatic exposure control (and Structured Illumination, if set).

The projector-specific controls are revealed by clicking on the SILab “Projector” program tab

Projector Output: manually changes the immediate projector output to crosses, a grid of dots, a single dot, a solid field of red/green/blue/black/white/ or an 8-bit grayscale value. Used for diagnostic purposes and not normally needed for routine imaging.

Image Registration: perform image registration for structured illumination imaging (see “3c. Multi-tile, SI Gel Imager mode” for more details)



The cutter-specific controls are revealed by clicking on the SILab “Cutter” program tab

Well #: the current well number

Wash Tip: perform one round of washing the cutter head

Select Spot: turn spot picking on or off

Cut and Deposit: Cut all currently picked spots and deposit them into respective wells

Image Cut Area: Deprecated. Do not use

Home Cutter: Deprecated. Do not use.

Align Cutter: Deprecated. Do not use.

3a. DIGE mode

To image with low dynamic range (most of the time, including SDS-PAGE minigels):

1. Load the gel by putting the gel cassette with the gel onto the stage. Lower the loading lid

2. Move the stage so the camera is overlooking your sample, by clicking on a tile and clicking “Go To Tile”, then moving the stage manually to the final position with Left/Down/Up/Right buttons

3. Set the gel origin (gel origin is the top left corner of your gel) by clicking “Set Gel Origin”

4. Set the image tile size (1x1 for a single tile, about 80% of a minigel, to 4x5, full coverage of a 2D gel)

5. Toggle the channels you want (Cy2/Cy3/Cy5) and set their exposure times in seconds

6. Click “Begin”. SILab asks for a filename, and when you click OK imaging will start.

3b. Structured Illumination mode

To image with high dynamic range (so-called Structured Illumination mode, SIGI mode or CPS mode) use the following steps:

3b1. Perform image registration:

1. Load the fluorescent uniform agarose gel by putting it at the center of the gel plate and loading the gel plate into the imager

2. Click “Gel Center” in the manual stage control

3. Click on “Projector” tab. Click “+ +” under “Projector Output”

4. Click “Register” under “Image Registration”

5. Follow SILab’s directions

6. Remove the agarose gel by clicking “Gel Load” in the manual stage control and taking the agarose gel off the gel plate

7. Load gel to be image by putting on the gel plate with destain.

3b2. Obtaining CPS images:

1. Load the gel by putting the gel cassette with the gel onto the stage. Lower the housing lid

2. Move the stage so the camera is overlooking your sample, by clicking on a tile and clicking “Go To Tile”, then moving the stage manually to the final position with Left/Down/Up/Right buttons

3. Set the gel origin (gel origin is the top left corner of your gel) by clicking “Set Gel Origin”

4. Set the image tile size (1x1 for a single tile, about 80% of a minigel, to 4x5, full coverage of a 2D gel)

5. Toggle the channels you want (Cy2/Cy3/Cy5) and click on SI to enable SI Gel Imager mode

6. Select the exposures you want (*eg.* 5 exposures starting with 0.125s with a 2x multiplier will image 0.125, 0.250, 0.5, 1 and 2 seconds for EACH tile and EACH channel).

6. Click “Begin”. SILab asks for a filename, and when you click OK imaging will start. If you have not done so, SILab will ask you to performed image registration (see 3b1). If the current image registration settings are good (nothing has been bumped), simply click on “Projector” tab, then click “Load” under “Image Registration”

WARNING: There is currently no graceful way of terminating a SI Gel Imager imaging session. Instead, force-quit the SILab program (Ctrl+Alt+Delete, start Task Manager, selecting SILab and pressing “End Task”). Then, you can manually reset the now-frozen stages (section 6a). Therefore it’s very important that you set up the imaging parameters correctly before starting !!!

4. SI Gel Imager shutdown sequence

1. Save all single-tile images (multi-tile images are automatically saved)

2. Press “Gel Load”, the gel plate will move the loading position.

3. Open the housing lid and remove the gel plate

4. If you have just finished cutting, remove the cutting head by unscrewing, removing it and storing it in a safe, sterile place (like a sterile 50mL conical tube)

5. Close the housing lid.

6. Turn off the projector by pressing its ON/OFF button twice. The projector power light will blink orange.

7. Press “QUIT” in SILab. SILab will confirm that you want to quit, press YES

8. The gel stage will move to its default position and the cutting head will lower.

9. Once the cutting head has finished lowering and the projector power light is solid orange, turn off the imager by pressing the master switch on the power strip.

5. Gel cutting

The gel cutting process involves taking a gel image, then selecting spots on the image by clicking on them in SILab and start a “Cut and Deposit” procedure. This extracts a small cylindrical plug of polyacrylamide gel containing a protein spot using the imager’s gel cutting head and depositing this plug into a 1.5mL test tube (SILab can deposit into a standard 96-well plate, but this is no longer done). First the cutting head pierces the gel, coring out the gel plug. Then, the plastic plunger inside the cutting head retracts, which creates suction that sucks the now-severed gel plug into the cutting head. The cutting head then moves over a test tube, lower and deposit the spot into the tube. While SILab allows cutting of a single plug at a time, or multiple spots in batch mode. It’s a good idea (albeit slower) to cut one plug at a time to verify

that the cut happened. SILab can cut anywhere on the gel plate, though it's a good idea to cut all the spots in the same area corresponding to a single imager tile before moving to another area in a distant part of the gel.

Successful cutting requires fluid to maintain suction. The gel being cut should be immersed in destain (40% methanol, 10% glacial acetic acid, 50% diH_2O) at all times to prevent drying and to provide this suction. Similarly, the gel plugs should also be deposited into fluid. In our typical collaboration with UPitt Proteomics Core, the gel plugs are deposited into 1% glacial acetic acid in HPLC-grade H_2O. For long term storage, however, all fluid should be removed from the tubes and the plugs stored at -20C or -80C until ready to be digested. Between cutting each plug, SILab also washes the cutter in the fluid, so cutting should not start until the test tubes containing fluid is also put into the gel plate.

WARNING: Take special care to not bump the cutting head while loading the gel plate into the imager. The tip of the cutting head is especially fragile and deforms easily. Once deformed, the gel head will rupture the plugs while cutting. To fix this, remove the cutting head from the shaft and reform it.

NOTE: It's very easy to contaminate the gel plugs with keratin from skin and hair. Wear gloves throughout the cutting process.

5a. Setting up gel plate and cutting head

1. Turn on imager and start SILab. Wait for the cutting motor to retract

2. If the cutting head is not attached, attach it by screwing the head onto the cutter

3. Determine the number of tubes needed for plugs: each 1.5mL Eppendorf normally holds one gel plug (two if the protein spot being cut is especially large). Be sure to use low-binding tubes. Fill up tubes with 1% glacial acetic acid in HPLC-grade water.

4. Put Eppendorf tubes into the plastic tube holder block. Leave tubes open and press down on the caps so they will not get caught on the cutter while the stage is moving.

5. Put plastic tube holder block containing tubes onto the gel plate.

5b. Imaging

Gel cutting requires acquiring a 1x1 tiled DIGE image of the cutting area. See section 3a.

5c. Performing the actual cut

1. With SILab running, click on the name of your 1x1 tiled DIGE image to select it

2. Click on the “Cutter” tab. Select “1.5mL tube”, spot #1 will be automatically selected (upper-left-most spot of the tube rack)

3. Click on “Select Spot”, the cursor will change into a magnifying glass

4. Hover over the image area and click to select the spot. A red circle with the spot number will show up on the image. To change the position of the spot click “Remove: Last” and select another spot.

5. Click “Cut and Deposit”. SILab will first wash the cutting head in the corresponding tube, move the stage to the spot position, cut, retract the head, move stage to the tube, deposit the plug, then retract the cutting head. SILab will then stop and the red circle on the image will turn blue, indicating the cut has completed.

6. Confirm the cut by taking a post-cut image: click on the 1x1 tiled image in the automatic stage control area and press “Goto Tile”. Click on the “Camera” tab and take a single-shot image by pressing “Acquire One”. The cut spot will show up a dark circle in the gel.

7. Click on the “Cutter” tab and click “Remove: All”.

8. Repeat selecting spots and cutting as necessary.

5d. Cleanup

1. Click “Gel Load”. The gel plate will be moved to the loading position

2. Remove the plastic tube holder block from the gel plate. Remove fluid from the Eppendorf tubes with a small syringe. The plugs can now be stored or shipped off for mass spec analysis.

3. Remove the gel plate from the imager

4. Shutdown the imager by quitting SILab

5e. Troubleshooting

Symptom:

Cause:

Solution:

6. Image analysis

Most common image analysis tasks are done in ImageJ (http://rsbweb.nih.gov/ij/) after the images have been acquired by SILab

6a. Making a 2-frame animated grayscale movie

1. Open both grayscale images in ImageJ

2. Open up the two TIFF files

3. Image > Stacks > Images to Stack, then click OK *

4. Once stack is assembled movie can be played by pressing the \ key (or Image > Stacks > Tools > Start Animation)

5. To save the movie: File > Save As > AVI. Change framerate to desired value, or simply accept the default options.

6b. Making a DIGE overlay false-color image

1. Open both grayscale images in ImageJ

2. Image > Lookup Tables > Green (for Cy3 image)

3. Image > Lookup Tables > Red (for Cy5 image)

4. Select Cy3 image, press Ctrl+A (or Edit > Selection > Select All)

5. Press Ctrl+C (or Edit > Copy)

6. Edit > Paste Control . Change drop-down to “Add”

7. Select Cy5 image. Image > Type > RGB Color

8. Press Ctrl+V (or Edit > Paste)

9. Save the resultant overlay image as JPEG, PNG, or TIFF

7. Uncommon tasks

7a. Resetting gel stage after a crash

Symptom: The gel stages are stuck

Cause: SILab (or Windows itself) had crashed, either through software failure or power outage

Solution: Log into the imager computer as ‘digeuser’. Reset the stage by double-clicking the SI Gel Imager stage reset program icon on the desktop.

WARNING: Make sure the paths of the stages are clear of objects (*eg*,. gel plates, tools) and body parts (*eg.* fingers, faces) As soon as you double click on the reset program’s icon the stages will move.

7b. Resetting camera after a crash

Symptom: The camera shutter doesn’t click, the images show up black, or SILab crashes while attempting to take images

Cause: Camera has become unresponsive and needs to be completely reset

Solution: Power down imager, shut down computer and restart everything from scratch

7c. Lubricating stage lead screws

Symptom: Metallic grinding/screeching noise while stage is moving

Cause: Dirt has accumulated on the lead screw enough to cause the stage motors to slip.

Solution: Clean dirt from leadscrew with a small pipette tip and put on lead screw grease

7d. Changing projector air filter

Symptom: A warning message is projected onto the gel

Cause: The projector air filter has exceeded its recommended life time.

Solution: Clean filter and reset filter timer. Refer to CP-X505 manual

7e. Changing projector lamps

Symptom: The projector output is dim or non-existent

Cause: The projector lamp has degraded or become non-functional

Solution: Change the projector lamp. Refer to CP-X505 manual

7f. Cleaning the light dump

Symptom: There is dust in images but the gels are clean

Cause: There is dust in the light dump under the gel plate

Solution: Clean room. Remove light dump, clean light dump with compressed air and replace.

7g. Realigning imager optics

Symptom: Images are incomplete, tilted or shifted

Cause: The filter wheel or one of the lens elements are out of alignment

Solution: Take test photos with grid paper and align optics accordingly

TIP: realignment will be much faster if one person manipulates the micrometer while another works with SILab.

REFERENCES

Adessi, C., Miege, C., Albrieux, C. and Rabilloud, T. (1997). Two-dimensional electrophoresis of membrane proteins: a current challenge for immobilized pH gradients. In *Electrophoresis*, vol. 18 (ed., pp. 127-35.

Aebersold, R. and Mann, M. (2003). Mass spectrometry-based proteomics. In *Nature*, vol. 422 (ed., pp. 198-207.

Amersham. (2002). Ettan DIGE User Manual. Uppsala.

Bertin, E. and Arnouts, S. (1996). SExtractor: Software for source extraction. *Astronomy and Astrophysics Supplement Series* **117**, 393-404.

Bertucci, E., Pilu, M. and Mirmehdi, M. (2003). Text selection by structured light marking for hand-held cameras. In *Document Analysis and Recognition, 2003. Proceedings. Seventh International Conference on*, (ed., pp. 555-559: IEEE.

Bjellqvist, B., Pasquali, C., Ravier, F., Sanchez, J. Ä. and Hochstrasser, D. (1993). A nonlinear wide-range immobilized pH gradient for two-dimensional electrophoresis and its definition in a relevant pH scale. *Electrophoresis* **14**, 1357-1365.

Bradski, G. (2000). The opencv library. *Doctor Dobbs Journal* **25**, 120-126.

Cagney, G., Amiri, S., Premawaradena, T., Lindo, M. and Emili, A. (2003). In silico proteome analysis to facilitate proteomics experiments using mass spectrometry. *Proteome Sci* **1**.

Cleland, W. (1964). Dithiothreitol, a new protective reagent for SH groups*. *Biochemistry* **3**, 480-482.

Cortizo, E., Yeras, A. M., Lepore, J. and Garavaglia, M. (2003). Application of the structured illumination method to study the topography of the sole of the foot during a walk. *Optics and lasers in engineering* **40**, 117-132.

Dale, G. and Latner, A. (1969). Isoelectric focusing of serum proteins in acrylamide gels followed by electrophoresis. *Clinica Chimica Acta* **24**, 61-68.

de St Groth, S. F., Webster, R. and Datyner, A. (1963). Two new staining procedures for quantitative estimation of proteins on electrophoretic strips. *Biochimica et biophysica acta* **71**, 377-391.

Erkmen, B. I. (2011). Computational ghost imaging for remote sensing applications. *Interplanet. Netw. Prog. Rep* **42**, 1-23.

Ernst, L. A., Gupta, R. K., Mujumdar, R. B. and Waggoner, A. S. (1989). Cyanine dye labeling reagents for sulfhydryl groups. *Cytometry* **10**, 3-10.

Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K. and Weissman, J. S. (2003). Global analysis of protein expression in yeast. In *Nature*, vol. 425 (ed., pp. 737-41.

Gong, L., Puri, M., Unlü, M., Young, M., Robertson, K., Viswanathan, S., Krishnaswamy, A., Dowd, S. R. and Minden, J. S. (2004). Drosophila ventral furrow morphogenesis: a proteomic analysis. In *Development*, vol. 131 (ed., pp. 643-56.

Gorg, A., Postel, W., Weser, J., Gunther, S., Strahler, J. R., Hanash, S. M. and Somerlot, L. (1987). Elimination of point streaking on silver stained two,Äêdimensional gels by addition of iodoacetamide to the equilibration buffer. *Electrophoresis* **8**, 122-124.

Görg, A., Weiss, W. and Dunn, M. J. (2004). Current two-dimensional electrophoresis technology for proteomics. In *Proteomics*, vol. 4 (ed., pp. 3665-85.

Herbert, B., Galvani, M., Hamdan, M., Olivieri, E., MacCarthy, J., Pedersen, S. and Righetti, P. G. (2001). Reduction and alkylation of proteins in preparation of two,Äêdimensional map analysis: Why, when, and how? *Electrophoresis* **22**, 2046-2057.

Hiraoka, Y., Sedat, J. W. and Agard, D. A. (1987). The use of a charge-coupled device for quantitative optical microscopy of biological structures. In *Science*, vol. 238 (ed., pp. 36-41.

Hiratsuka, A., Kinoshita, H., Maruo, Y., Takahashi, K., Akutsu, S., Hayashida, C., Sakairi, K., Usui, K., Shiseki, K. and Inamochi, H. (2007). Fully automated two-dimensional electrophoresis system for high-throughput protein analysis. *Analytical chemistry* **79**, 5730-5739.

Hunter, A. Ä., Laursen, S. L. and Seymour, E. (2007). Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development. *Science Education* **91**, 36-74.

Kainz, F., Bogart, R. and Hess, D. (2004). The OpenEXR image file format. *GPU Gems: Programming Techniques, Tips and Tricks for Real-Time Graphics, R. Fernando, Ed. Pearson Higher Education*.

Kang, Y., Techanukul, T., Mantalaris, A. and Nagy, J. M. (2009a). Comparison of three commercially available DIGE analysis software packages: minimal user intervention in gel-based proteomics. In *J Proteome Res*, vol. 8 (ed., pp. 1077-84.

Kang, Y., Techanukul, T., Mantalaris, A. and Nagy, J. M. (2009b). Comparison of three commercially available DIGE analysis software packages: minimal user intervention in gel-based proteomics. *Journal of proteome research* **8**, 1077-1084.

Kerenyi, L. and Gallyas, F. (1973). úber Probleme der quantitativen Auswertung der mit physikalischer Entwicklung versilberten Agarelektrophoretogramme. *Clinica Chimica Acta* **47**, 425-436.

Klose, J. (1975). Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. *Humangenetik* **26**, 231-243.

Kwak, Y. and MacDonald, L. (2000). Characterisation of a desktop LCD projector. *Displays* **21**, 179-194.

Lange, V., Picotti, P., Domon, B. and Aebersold, R. (2008). Selected reaction monitoring for quantitative proteomics: a tutorial. In *Molecular Systems Biology*, vol. 4 (ed.

Macko, V. and Stegemann, H. (1969). Mapping of potato proteins by combined electrofocusing and electrophoresis identification of varieties. *Hoppe-Seyler's Zeitschrift fur Physiologische Chemie* **350**, 917-919.

Marouga, R., David, S. and Hawkins, E. (2005). The development of the DIGE system: 2D fluorescence difference gel analysis technology. In *Anal Bioanal Chem*, vol. 382 (ed., pp. 669-78.

Minden, J. (2007). Comparative proteomics and difference gel electrophoresis. In *BioTechniques*, vol. 43 (ed., pp. 739, 741, 743 passim.

Minden, J. S., Dowd, S. R., Meyer, H. E. and Stuhler, K. (2009). Difference gel electrophoresis. *Electrophoresis* **30 Suppl 1**, S156-61.

Molloy, M. P., Herbert, B. R., Walsh, B. J., Tyler, M. I., Traini, M., Sanchez, J. C., Hochstrasser, D. F., Williams, K. L. and Gooley, A. A. (1998). Extraction of membrane proteins by differential solubilization for separation using two-dimensional gel electrophoresis. In *Electrophoresis*, vol. 19 (ed., pp. 837-44.

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. In *J Biol Chem*, vol. 250 (ed., pp. 4007-21.

Pence, W. D., Chiappetti, L., Page, C. G., Shaw, R. and Stobie, E. (2010). Definition of the flexible image transport system (fits), version 3.0. *A&A* **524**, A42.

Schrimpf, S. P., Weiss, M., Reiter, L., Ahrens, C. H., Jovanovic, M., Malmström, J., Brunner, E., Mohanty, S., Lercher, M. J., Hunziker, P. E. et al. (2009). Comparative functional analysis of the Caenorhabditis elegans and Drosophila melanogaster proteomes. In *Plos Biol*, vol. 7 (ed., pp. e48.

Sellers, K., Miecznikowski, J., Viswanathan, S., Minden, J. and Eddy, W. (2007). Lights, Camera, Action! Systematic variation in 2-D difference gel electrophoresis images. In *Electrophoresis*, vol. 28 (ed.

Shimomura, O., Johnson, F. H. and Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *Journal of cellular and comparative physiology* **59**, 223-239.

Speers, A. E. and Wu, C. C. (2007). Proteomics of integral membrane proteins--theory and application. In *Chem Rev*, vol. 107 (ed., pp. 3687-714.

Srinivasan, V. and Lumia, R. (1988). A 3D vision system for robot guidance with structured sine wave illumination. In *Intelligent Control, 1988. Proceedings., IEEE International Symposium on*, (ed., pp. 196-200: IEEE.

Stahl-Zeng, J., Lange, V., Ossola, R., Eckhardt, K., Krek, W., Aebersold, R. and Domon, B. (2007). High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. In *Mol Cell Proteomics*, vol. 6 (ed., pp. 1809-17.

Unlu, M., Morgan, M. and Minden, J. (1997). Difference gel electrophoresis. A single gel method for detecting changes in protein extracts. In *Electrophoresis*, vol. 18 (ed.

Voss, T. and Haberl, P. (2000). Observations on the reproducibility and matching efficiency of two-dimensional electrophoresis gels: consequences for comprehensive data analysis. In *Electrophoresis*, vol. 21 (ed., pp. 3345-50.

Walker, J. M. (1984). Gradient SDS polyacrylamide gel electrophoresis. In *Proteins*, pp. 57-61: Springer.

Whiteside, U., Pantelone, D. W., Hunter-Reel, D., Eland, J., Kleiber, B. and Larimer, M. (2007). Initial Suggestions for Supervising and Mentoring Undergraduate Research Assistants at Large Research Universities. *International Journal of Teaching & Learning in Higher Education* **19**.

Zabel, C. and Klose, J. (2009). High-resolution large-gel 2DE. *Methods Mol Biol* **519**, 311-38.

Zabel, C., Mao, L., Woodman, B., Rohe, M., Wacker, M. A., Klare, Y., Koppelstatter, A., Nebrich, G., Klein, O., Grams, S. et al. (2009). A large number of protein expression changes occur early in life and precede phenotype onset in a mouse model for huntington disease. *Mol Cell Proteomics* **8**, 720-34.

Zhou, S., Bailey, M. J., Dunn, M. J., Preedy, V. R. and Emery, P. W. (2005). A quantitative investigation into the losses of proteins at different stages of a two-dimensional gel electrophoresis procedure. In *Proteomics*, vol. 5 (ed., pp. 2739-2747.

Zitova, B. and Flusser, J. (2003). Image registration methods: a survey. In *Image and vision computing*, vol. 21 (ed., pp. 977-1000.