Phu T. Van
Victor Bass
Dan Shiwarski
Frederick Lanni
Jonathan Minden

Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA, USA

Received March 11, 2014
Revised May 14, 2014
Accepted June 10, 2014

Research Article

High dynamic range proteome imaging with the structured illumination gel imager

A current challenge for proteomics is detecting proteins over the large concentration ranges found in complex biological samples such as whole-cell extracts. Currently, no unbiased, whole-proteome analysis scheme is capable of detecting the full range of cellular proteins. This is due in part to the limited dynamic range of the detectors used to sense proteins or peptides. We present a new technology, structured illumination (SI) gel imager, which detects fluorescently labeled proteins in electrophoretic gels over a 1 000 000-fold concentration range. SI uses computer-generated masks to attenuate the illumination of highly abundant proteins, allowing for long exposures of low-abundance proteins, thus avoiding detector saturation. A series of progressively masked gel images are assembled into a single, very high dynamic range image. We demonstrate that the SI imager can detect proteins over a concentration range of approximately 1 000 000-fold, making it a useful tool for comprehensive, unbiased proteome-wide surveys.

Keywords:
2DE gels / Fluorescence / Proteomics / Structured illumination
DOI 10.1002/elps.201400126

1 Introduction

A major challenge for comparative proteomics is the large range of protein concentration in complex biological samples. 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 [1]. Mammalian serum or plasma samples have even higher protein concentration ranges, reported to be in the millions- to billions-fold range [2]. These large concentration ranges pose an important hurdle for unbiased, proteome-wide comparisons that search for changes in protein concentration and posttranslational 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 MS for protein quantification by the detection of peptides arising from enzymatic digestion of protein samples [3]. This commonly used method employs LC followed by tandem MS (LC-MS/MS) to detect peptides over a ~1000-fold concentration range [4]. 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 [5]. Therefore, SRM cannot be used for unbiased, comparative proteome studies.

The protein-centric approach divides this comparative proteomics method into two main tasks: (i) separation of intact proteins by 2DE followed by quantitative protein detection and (ii) identifying proteins of interest by MS. DIGE is a commonly used method for comparing different fluorescently tagged proteome preparations. Standard 2DE gels are capable of resolving ~2000 protein spots per gel; large-format 2DE gels can resolve more than 10 000 spots [6]. 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 [7]. 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 posttranslational modifications.

Quantification of protein spots 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 [8]. Since the samples are subjected to the

Correspondence: Dr. Jonathan Minden, Department of Biological Sciences, Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, PA 15213, USA
E-mail: minden@cmu.edu
Fax: +1-412-268-7129

Abbreviations: ADH, alcohol dehydrogenase; **CPS**, counts per second; **DEE**, *Drosophila* embryo extract; **SI**, structured illumination; **SRM**, selected reaction monitoring

Colour Online: See the article online to view Figs. 1, 2,6 and 8 in colour.

same electrophoretic forces simultaneously, intergel 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 [9]. Previous studies estimated that full-field and laser scanning imaging of 2D-DIGE gels has a protein detection range of about 10 000-fold, defined as the ratio between highest and lowest fluorescent intensity of detected spots [10]. While promising, this performance still falls short of the dynamic range required for whole-cell proteome studies.

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 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 protein concentration range [11]. 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 which means that different parts of the imaging area are illuminated with different amounts of light. SI has been extensively studied using different techniques toward different goals, with applications in robot guidance [12], remote sensing [13, 14], three-dimensional surface scanning [14], and document processing [15]. There are also notable SI applications in microscopy in which SI is employed for axial and transverse super-resolution (Zeiss Apotome2, Leica OptiGrid). To date, no SI application involves the imaging of electrophoretic gels. In the application of SI 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 (unmasked) pixel intensity values remain reliable, allowing for the detection of proteins over a 1 000 000-fold protein concentration range in a single field of view of an SDS-PAGE gel and significant increase in the number of protein spots detected in a 2DE gel containing whole extracts.

2 Materials and methods

2.1 The SI imager

The central principle of the SI 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 predetermined 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 $2N$ exposure is captured. The new image is used to compute the next masked image that is projected onto the gel and a $4N$ exposure is recorded. Once a pixel is masked, it remains masked throughout the imaging series. Typically, this cycle is repeated six to eight times, with a maximum exposure time of 64 s. 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. 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. 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. 1C). In Image A, obtaining a nonsaturated 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, such as the spot indicated by the green arrow, can be boosted to high dynamic range.

2.2 Construction of the SI gel imager

The main components of the SI gel imager are illustrated in Fig. 2A. An auditorium-grade 3-LCD projector, with its standard lens removed, was used as the SI light source (NEC

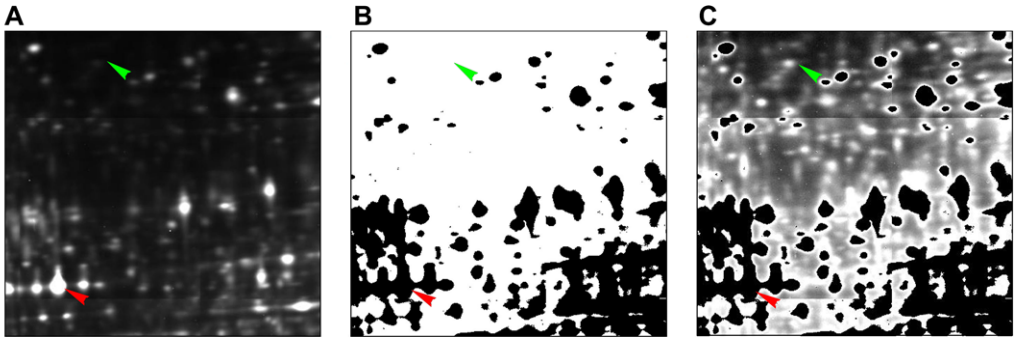

Figure 1. Structured illumination improves quality of fluorescent gel images. (A) Fluorescence image of a typical 2DE protein gel. Abundant proteins appear as bright spots (red arrow) while low-abundance proteins are not visible (green arrow). (B) A binary mask generated from (A): pixels above a certain 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.

Display of America, Itasca, IL). For this model projector, light is combined from three single-color LCD arrays into a single beam of near-full spectrum white light for output. Light emitted from the horizontally mounted projector was focused through a pair of achromatic lenses to form an image onto the gel ($f = 200$ mm, Edmund Optics, Barrington, NJ), then passed through fluorescence excitation band-pass filters mounted in a motorized filter wheel (CVI, Albuquerque, NM) and redirected onto the gel via a multibandpass 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, antireflection 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). For imaging fluorescent samples, the 2″-diameter optical filters for Cy3 (excitation: 545 ± 12.5 nm and emission: 590 ± 15 nm) and Cy5 (excitation: 635 ± 12.5 nm and emission: 695 ± 27.5 nm) were from Chroma Technology (Bellows Falls, VT), though any excitation and emission filter sets could be used depending on the fluorophores being employed. 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, Ft. Wayne, IN).

The projector had a native resolution of 1024 × 768 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 × 1340 pixels. For the SI gel imager, the CCD's acquisition field was cropped to the central 1024 × 1024 pixels of the CCD and binned (4 × 4) to produce a 256 × 256 pixel image. These CCD adjustments ensure projector oversampling and increase S/N ratio of the camera image. To minimize projector spatial nonuniformity, 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 [16]. The gel plate accommodated 2DE gels up to 250 × 200 mm. The camera field-of-view was 42 × 42 mm, 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 1 μm resolution (NEAT, Beverly, MA).

The SI gel imager is controlled by a dedicated computer workstation (Dell, 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.

2.3 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 field-of-view, a process called "image registration." Many image registration methods exist to satisfy different constraints on registration speed, accuracy, resistance

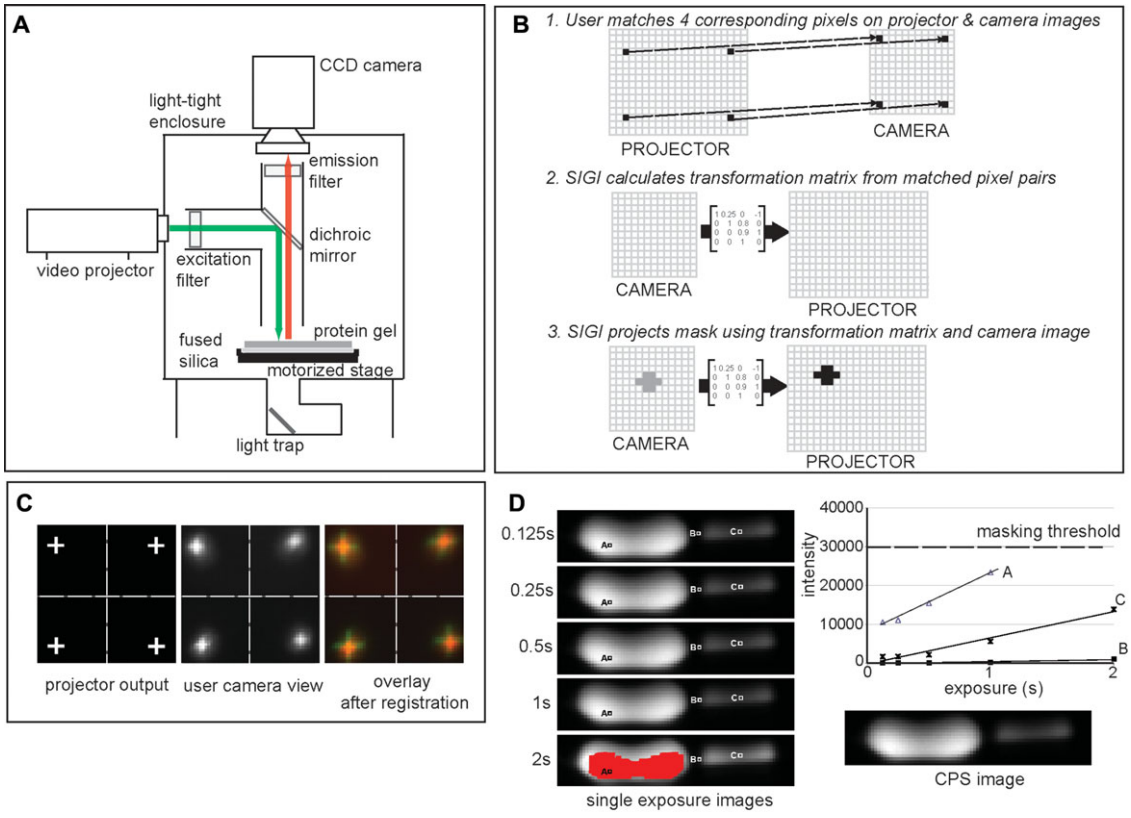

Figure 2. Components and operation of the SI gel imager. (A) Diagram of the SI gel imager's optical configuration. (B) Workflow of the SI gel imager's image registration process. (C) Shown here are the image registration results where subimages of the four corners of imager's field-of-view are displayed. 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.

to image noise, and automation [17]. We require pixel-level alignment for the entire camera field-of-view with a relatively simple optical configuration (rigid planar camera and projector surfaces, and 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. 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 (5 × 5 pixel "+" signs) that are positioned at the four corners of the camera field-of-view onto a uniformly fluorescent target—a 1.5 mm thick, agarose slab gel containing Cy3-labeled protein (serum albumin) (Fig. 2C, left panel). An enlarged version of the resultant camera image is displayed on SILab viewer (Fig. 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 as:

$$[\mathrm{CCDx}, \mathrm{CCDy}] = \mathrm{M} \times [\mathrm{Proj}_{x,} \mathrm{Proj}_{y}] \quad (1)$$

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 tool kit [18]. OpenCV's getPerspectiveTransform() function takes Proj_x, Proj_y from the projector image cross-hair 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. 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. Unsupervised methods were also tested, but performed poorly in the presence of slight defects in the target gel, while users can visually pick the crosses quickly and correctly regardless of minor target gel defects. 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 min.

2.4 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. It has been established that CCD cameras have highly linear responses up to saturation [19,20] making them suitable for quantitative measurements. Consistent with these reports, the 16-bit CCD camera used in the SI gel imager displayed a response that was linear with all exposures up to about 90% of saturation. Given this linear response and the fact that all exposure times are known in an SI gel imager image series, 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:

$$Y(c,t) = c_0 + c_1 t \tag{2}$$

where Y is observed intensity in counts, t is exposure time in seconds, and c_0 and c_1 are the best-fit offset and slope. Using the GNU Scientific Library (Free Software Foundation, Boston, MA) mathematical tool kit, 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 as:

$$\%\varepsilon = \frac{\left(\sum_{i=1}^{n} \sqrt{\frac{(y-\hat{y})^2}{n}}\right) \times 100}{\mathrm{CPS}} \tag{3}$$

where %ε is the pixel's percent residual, y is the observed intensity, $\hat{y}$ is the interpolated intensity, and 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 10× 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, Uppsala, Sweden). A series of five cycles of masked SI images were taken with 0.125, 0.250, 0.5, 1, and 2 s exposures (Fig. 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.

A simple file format for CPS images was devised to streamline data storage and analysis. The header of the file records image dimensions and fluorophore wavelengths. The data portion stores CPS intensity values as 32-bit floating-point arrays, which is well beyond the dynamic range of current gel imagers. This raw image format is readily handled by common image analysis software packages, such as ImageJ.

2.5 Postcapture image processing

Since complete camera coverage of a typical 2DE gel requires 4 × 5 tiles (1024 × 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 [21]. When applied to SI images, this model appears as:

$$\pi_{x,y} = (Y_{x,y} - \rho_{x,y})/\alpha_{x,y} \tag{4}$$

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 BSA (New England Biolabs) was dissolved in the polyacrylamide gel solution before polymerization. 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 in a petri dish in the dark with shaking for 1 h at 37°C to reduce any disulfide bonds in the BSA, and 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 min and imaged. To minimize artifacts from fluorescent dust particles or local inhomogeneities in the gel, the median image from nine 1-s 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 above equation.

2.6 Evaluating imager detection capability

To determine the imager's practical dynamic range, we prepared dilutions of purified protein samples in a typical SDS-PAGE minigel (which will be referred to as a 1D gel). We generated a tenfold 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 tenfold increase in fluorescent labeling relative to manufacturer's recommended minimal labeling. The stock solutions were diluted with an equal volume of 2× 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 min. The tenfold 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 2× Laemmli buffer, 2.5 μL βME, 10 μL of unlabeled 10 mg/mL BSA as carrier protein, and 40 μL of 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 min and 5 μL of each sample was loaded into consecutive lanes of a 12% polyacrylamide minigel 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 min. 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–1 μg, while the box for the 10 μg lane was increased

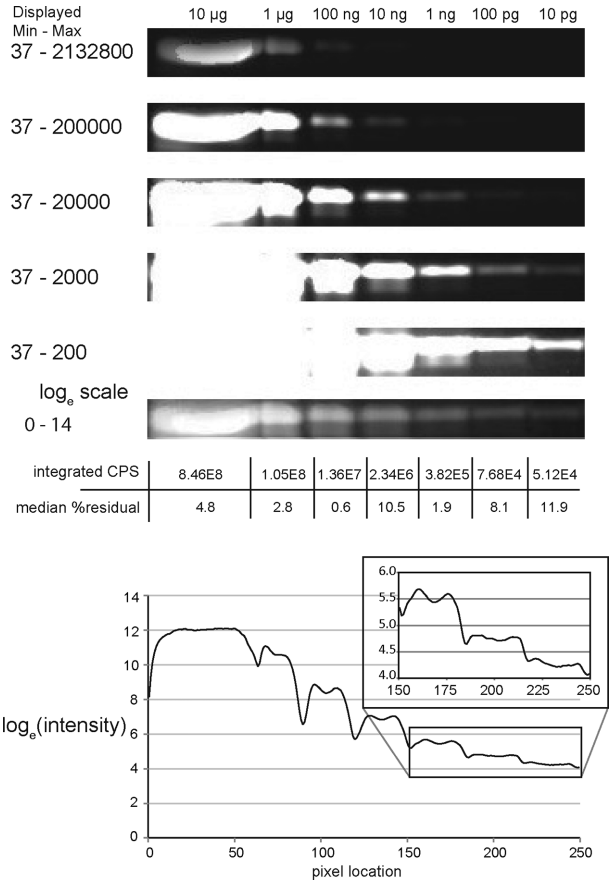

Figure 3. The SI gel imager detects protein over a 1 000 000-fold concentration range. Seven tenfold dilutions of purified, Cy3-ADH from 10 μg to 10 pg were electrophoresed in a 1D SDS-PAGE gel and imaged with the SI gel imager to generate a CPS image. Successively decreasing the displayed maximum pixel intensity reveals progressively dimmer bands (first to fifth 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.

in size to accommodate the high amount of protein loaded in this lane (Fig. 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 [22]) was labeled with 1 nmol of Cy3-NHS and incubated on ice for 20 min; 1 μL of quencher (5 M methylamine-HCl and 10 mM HEPES (pH 8.0)) was added and the reactions were then incubated for 2 h before being run on 18 cm 3-10NL IEF strips—this labeling extent will be referred to as 1×-labeled (dyes and IEF strips: GE Healthcare) for 32 000 Vh. Second-dimension SDS-PAGE was done in 18 × 20 cm gradient gels (10–15% polyacrylamide), electrophoresed in Tris-glycine-SDS running buffer at 300 V until the tracking dye ran off the bottom of the gel (5–6 h). 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 10×-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, floating-point CPS image generated from 12 SI 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 2 × 2 tiles (512 × 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. CPS images were converted from 32-bit raw, floating-point, pixel arrays into 16-bit, unsigned integer, TIFF images in ImageJ as required for DeCyder analysis. The 2DE DEE images were also analyzed using SourceExtractor (abbreviated as "SExtractor" [23]), an open-source astronomy software package that detects point sources in dim images that we had used to analyze 2DE gels previously [10]. Conceived as a galaxy counting tool, SExtractor subtracts background signal from input images to look for Gaussian point sources that satisfied oval shape constraints. CPS images were converted from 32-bit raw, floating-point, pixel arrays into formatted 32-bit, floating-point, Flexible Image Transport System images in ImageJ as required for SExtractor analysis. Though fundamentally different in their approaches, both DeCyder and SExtractor showed similar trends in results, as discussed later. For quantification in ImageJ, unprocessed CPS images were used, and data analysis was done in Matlab (Mathworks, Natick, MA) and Excel (Microsoft, Redmond, WA).

3 Results and discussion

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

3.1 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. 3). 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 10× labeling). We did not encounter any protein solubility problems associated with the increased dye labeling. Only a very modest amount of protein spot spreading along the molecular weight axis in 2DE gels was observed, which was anticipated since the tenfold increase in labeling was expected to increase the number of proteins with two or more dye molecules to a very minor extent, while significantly increasing the pool of singly labeled proteins. To test the extent of photobleaching, an agarose gel containing a uniform distribution of Cy3-NHS-labeled BSA was illuminated with Cy3 excitation light from the SI gel imager's video projector for 16 min and showed only a 1.8% reduction in fluorescence intensity. Thus, typical SI gel imager imaging sessions, in which individual tiles are illuminated for much less than 16 min, are not expected to suffer from photobleaching. Progressively decreasing the image's displayed maximum pixel intensity value allows one to see the dim protein bands (Fig. 3, first to fifth 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. 3, sixth 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 of protein (inset images). The CPS image had a maximum intensity of 2.13 million counts, far exceeding the 65 536 count upper limit allowed by single-exposure 16-bit images. The integrated intensity of the 10 pg band was 5.12×10^4 CPS. 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 10×-labeling. The same million-fold dilution experiment with ADH labeled at the recommended Cy3-NHS dye concentration (1× labeling) only provided a detection range of the five most abundant lanes in the CPS image, a ~10 000-fold dynamic range (Fig. 4). Similarly high dynamic range images were captured for Cy5-NHS-labeled ADH, indicating that the SI gel imager performance is independent of the fluorophore used, allowing for the detection of low-abundance protein differences using DIGE labeling schemes (data not shown).

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,

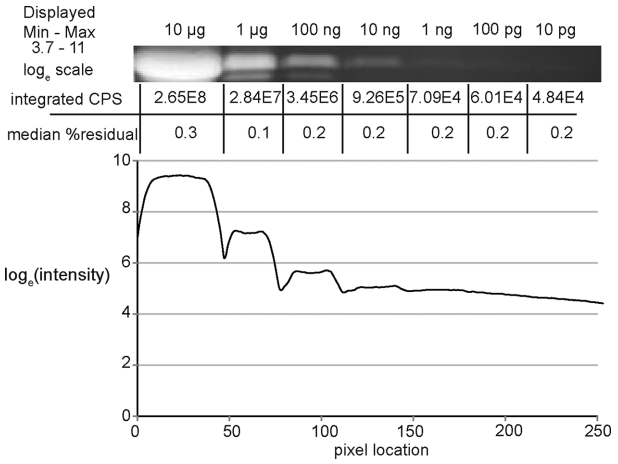

Figure 4. Million-fold dilution series (10 µg–10 pg) of 1×-labeled Cy3-ADH. Only the most abundant five lanes (10 µg–1 ng) are visible in the CPS image.

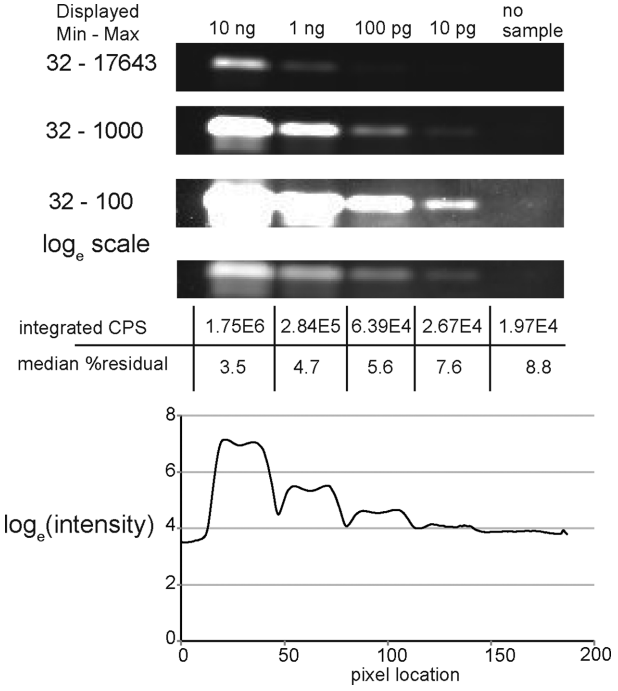

Figure 5. A thousand-fold dilution series (10 ng–10 pg) of ADH labeled with 1X Cy3-NHS. The last four samples of the spin-dialyzed, 10×-labeled Cy3-ADH dilution series from Fig. 3 were imaged with the SI gel imager. This image clearly shows the presence of the 10 pg sample relative to the adjacent empty lane.

this slight contamination made quantification of the low abundance difficult. To confirm the detection and quantification of the 10 pg band, we ran a 1000-fold dilution series on a 1D gel using the lowest four spin-dialyzed samples (10 ng–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. 5). The residual for the 10 pg band, which had an integrated fluorescence signal of 2.67×10^4 CPS, was 7.6%, making it distinguishable from the integrated intensity of 1.97×10^4 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 protein detection range of at least 1 000 000-fold in 1D gels loaded with 10×-labeled proteins. It is important to note that this experiment was intended to test the LOD for the SI gel imager. Inspection of the integrated fluorescence intensity of individual bands was only roughly linear and the slope did not reflect the tenfold concentration difference between lanes. We do not entirely understand this lack of direct correspondence between loaded protein and integrated fluorescence signal. This is mostly likely due to a combination of gel crowding and modest dye:dye quenching effects at higher protein concentrations. Regardless of the explanation, the SI gel imager is capable of detecting proteins over a million-fold concentration range and the output from this dilution series can potentially be used to generate a calibration curve relating protein amount to fluorescence intensity.

3.2 High dynamic range detection of complex protein mixtures in 2D gels

To assess the efficacy of the SI gel imager for the analysis of cellular proteomes, a series of 2DE gels containing Cy3-NHS 1×-labeled *Drosophila* embryo extract was analyzed. The goal was to compare the detection range and reliability of single-exposure (unmasked) images to SI gel imager-generated CPS images. These images were analyzed using several image analysis packages: (i) DeCyder (version 5.0), a commercial, 2DE gel analysis software package from Amersham BioSciences (now GE Healthcare Life Sciences), (ii) Source Extractor (SExtractor, version 2.8), an open-source astronomy analysis application that was adapted to analyze 2DE gel images, and (iii) 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 × 2 tiled region of a full 2DE gel (Fig. 6A), DeCyder detected an 8% (336/311) increase in raw spot count between the single-exposure image and the CPS image (Table 1). Since 2DE gels of fluorescent proteins tend to be noisy, unsupervised, raw spot counts by

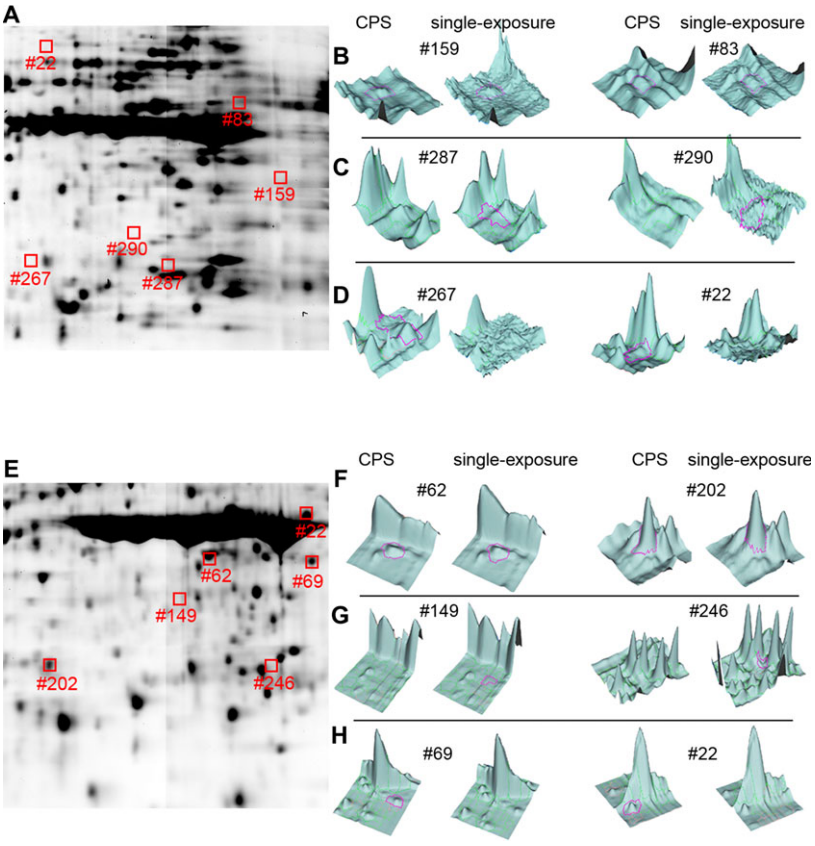

Figure 6. Comparison of DeCyder spot detection of CPS and single-exposure images. (A) A CPS image of a 1×-labeled Cy3-DEE sample resolved on a 2DE gel. DeCyder was used to detect protein spots in CPS and single-exposure images of this 2DE gel. The numbered protein spots indicate a set particular spots chosen to highlight the differences between the CPS and single-exposure images. (B) 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) 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) 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) A CPS image of a 10×-labeled Cy3-DEE sample resolved on a 2DE gel. (F) A pair of positively detected protein spots in the CPS and single-exposure images of the gel shown in (E). (G) 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) 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).

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

	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	2360
Brightest protein spot (flux)	27 800 000	26 900 000
Summed yolk region (flux)	102 000 000	90 100 000
Flux ratio (brightest/dimmest)	2300	11 400
Dynamic range (summed yolk/dimmest)	8500	38 100

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. 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 lead the rejection of 53 putative protein spots, a 16% reduction (Table 1). 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).

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. 6 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. 6B). However, image noise caused DeCyder to falsely assign spots to the "shoulder" of an existing spot (Fig. 6C, left panel), or to background areas (Fig. 6C, 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. 6D, 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 2400, whereas the flux ratio for the CPS image was 12 000 (Table 1). Therefore, the dynamic range of 1×-labeled

Drosophila embryo extract according to the CPS image was fivefold 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 posttranslationally 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 an LC-MS/MS "shotgun proteomics" approach and reported a dynamic range of approximately 10 000-fold [24]. 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 MS. 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 [22]). Comparing the flux of dimmest protein spot to the summed yolk protein signal increased the dynamic range to 8480 and 38 121 for the single-exposure and CPS images, respectively. This represents a nearly fourfold increase in dynamic range. It also indicates a fourfold improvement in detected concentration range in *Drosophila* embryos.

3.3 The effect of excess fluorescent labeling on protein spot detection

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

DeCyder spot extraction for a 10×-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 2). 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 10×-labeled sample was similar to that of the 1×-labeled sample. Surprisingly, the dynamic range of the CPS image for the 10×-labeled sample was slightly lower than the single-exposure image of the same gel with the 10×-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 1×- to 10×-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 10×-labeled DEE gel. Despite the increase in fluorescence signal, the DeCyder measured flux emanating from the yolk protein spots did not increase tenfold, as expected. Instead, when compared to the same area in 1×-labeled DEE 2DE gels, the 10×-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 10×-labeled yolk proteins spots appear larger mostly in the second dimension, not in the first dimension; because of the increased labeling that 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. 7). This molecular crowding may explain why the 10×-labeled yolk protein peaks did not have the expected Gaussian shape, but appeared to be somewhat plateaued (Fig. 7).

Table 2. DeCyder spot detection and quantification of a 2DE gel containing 10×-labeled DEE

	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)	2750	2350
Dynamic range (summed yolk/dimmest)	15 420	13 700

3.4 High-abundance protein flux measurement errors

The discrepancy in flux ratio between 1×- and 10×-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 1×-labeled sample to the dimmest spot detected in the 10×-labeled sample showed an approximate tenfold 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 (Tables 1 vs. 2). Comparing the brightest individual spot and the summed yolk region of the 1×- and 10×-labeled samples revealed only a 1.6- to 2.7-fold flux increase, respectively (Tables 1 vs. 2). While low-abundance, well-separated protein spots roughly reflected the tenfold increase in labeling extent, high-abundance proteins failed to reflect the expected tenfold increase in labeling. In DeCyder, the background value for any spot segment is taken as the lowest tenth-percentile value of the spot border, not the gel's

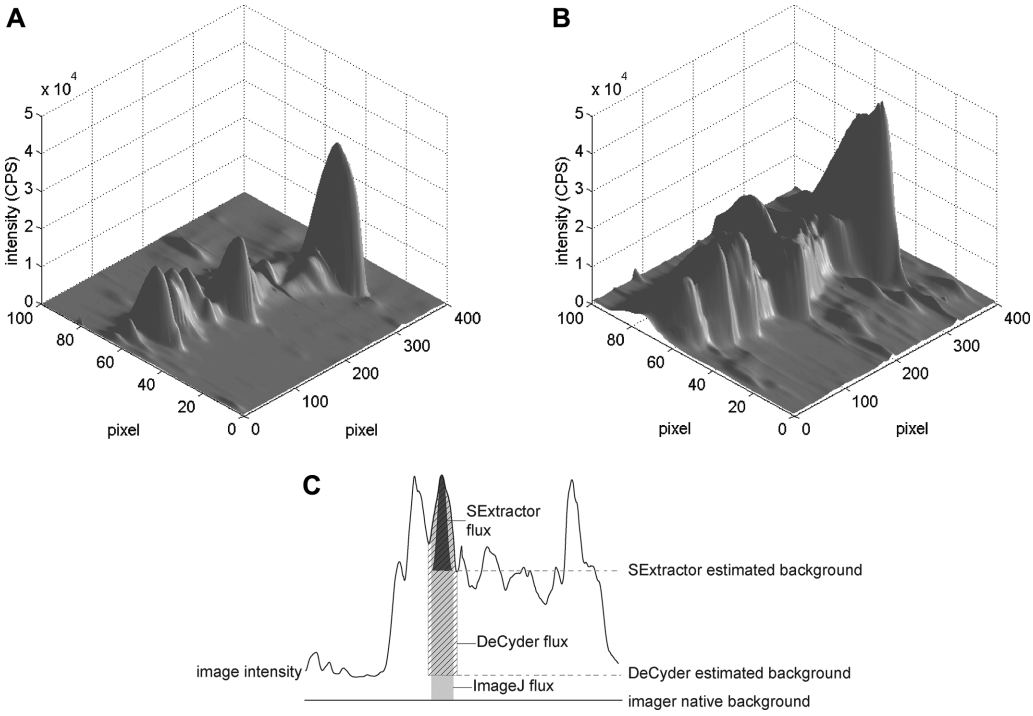

Figure 7. Surface plots of the yolk protein regions of DEE 2DE gels. (A) A Matlab-generated surface plot from the CPS image of the yolk protein region of a 1×-labeled DEE 2DE gel. (B) The same region of a 10×-labeled DEE sample. (C) Intensity cross-section of the yolk protein region of a DEE gel. To 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.

native background signal [25]. 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 3 and 4). 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 [23]. Since it was conceived as a galaxy detection tool, SExtractor assumes that the input image consists of a dark field background overlaid with overlapping, roughly Gaussian peaks corresponding approximately to elliptical galaxies. Given the high dynamic range of star field images, SExtractor's operations are 32-bit throughout, preserving raw pixel intensities without scaling. SExtractor attempts to first estimate the background pixel intensity, then given different, user-controlled parameters, separate the remaining pixels into

Table 3. SExtractor/ImageJ spot detection and quantification of a 2DE gel containing 1×-labeled DEE

	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	3790
Dynamic range (summed yolk spots/dimmest)	18 500	78 400

Table 4. SExtractor/ImageJ spot detection and quantification of a 2DE gel containing 10×-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	1470
Dynamic range (summed yolk spots/dimmest)	113 000	71 100

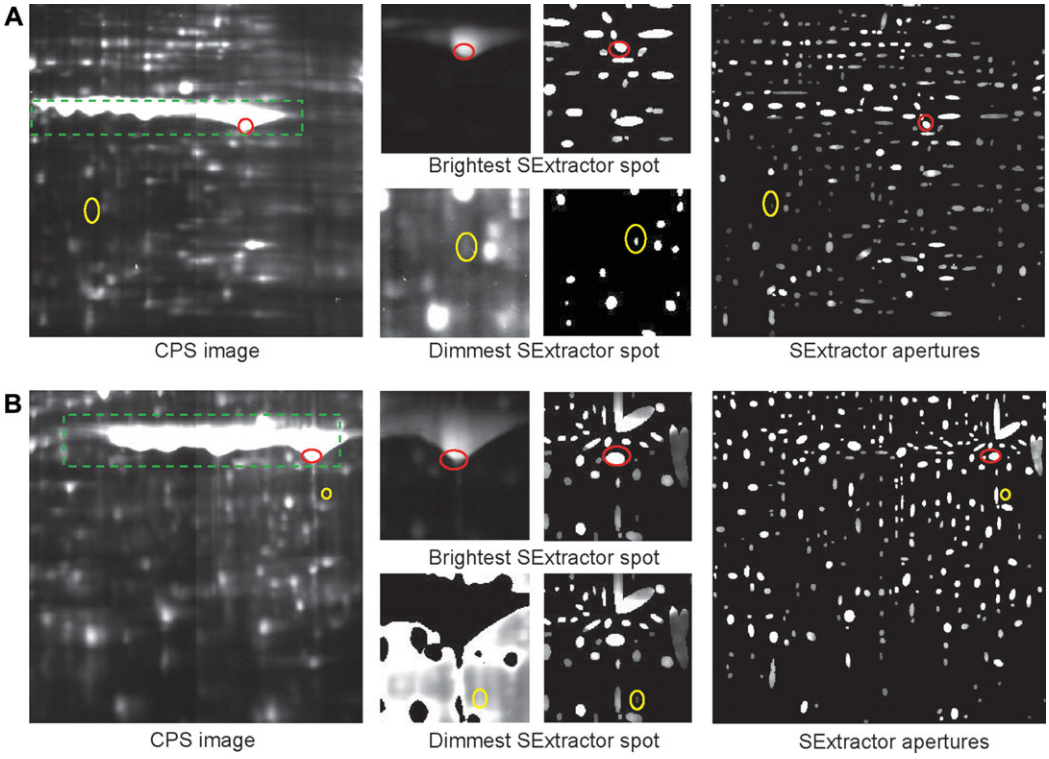

Figure 8. SExtractor analysis of CPS images. The same DEE 2DE gels analyzed by DeCyder in Fig. 4 were analyzed using SExtractor. (A) The 1×-labeled DEE 2DE gel. The leftmost panel displays the CPS image, whereas the right-most panel shows SExtractor's detected apertures. The upper pair of smaller middle 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) The 10×-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.

objects resembling 2D Gaussian peaks called apertures in a deblending process. Deblending is controlled by parameters that include spot size range, spot packing, and background estimation, which we set to values appropriate for finding spots on 2DE gels (e.g. apertures should be larger than five pixels on the minor axis). 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 preset 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. Examples are illustrated in Fig. 8: the brightest SExtractor spot comprised only a small fraction of the yolk spot (Fig. 8A 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 1×- and 10×-labeled samples (Fig. 8A 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. 8A, 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. 8B, 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 [26], ImageJ was used to manually draw a rectangular selection box around the yolk protein area (Fig. 8, green, dashed-line rectangles) of CPS images, which were kept as 32-bit raw pixel arrays to avoid loss of precision. 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 3 and 4). The vast majority of proteins within the yolk protein box are known to be yolk protein, therefore it is reasonable to perform a simple integration of the detected spots in 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 1×- and 10×-labeled samples, respectively (Tables 3 and 4). 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 1×- and 10×-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 1×-labeled sample was not sufficient to detect the full concentration range of proteins in a sample, then 10× labeling would be expected to reveal a wider concentration range. Since the detected protein concentration range was roughly similar between 1× and 10× labeling, one can assume that the SI gel imager detected the full range of proteins in these embryo extracts.

4 Concluding remarks

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 LOD of the SI gel imager is approximately 10 pg of 10×-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 antireflection-coated, fused-silica plates upon which the gels were mounted. While the SI gel imager yielded improved results employing the common practice of imaging 2D-DIGE gels within low-fluorescence glass plates, the background emanating from standard low-fluorescence plates was more than four times that of the fused-silica plates in the Cy-3 channel. We found that the BBAR coating was not stable over long periods of use. A more rugged antireflection 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 in influencing the SI gel imager's LOD. The LOD for 1×-labeled ADH was about 1 ng, whereas 10×-labeled ADH allowed for the detection of as little as 10 pg of ADH. In 1× labeling, approximately 5% of all protein molecules carry a single dye molecule, the rest are unlabeled. In 10× labeling, about half of all 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. 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. While this report focuses on covalently attached fluorescent dyes, it will be interesting to examine the SI gel imager's LOD for fluorescent protein stains such as Flamingo, Krypton, and SyproRuby.

With respect to imaging, it is possible to extend imaging dynamic range by simply excluding saturated pixels from a series of exposures and assembling intensity information into a composite image. However, compared to this approach, SI has two benefits: one with respect to the exposure of fluorescent proteins within the gel matrix and one with respect to CCD behavior. First, long exposures of areas containing high-abundance proteins, which would ordinarily saturate the detector, also emit enough scattered light to degrade signals from nearby, low-abundance proteins. Masking limits this scattered light effect, increasing the detectability of nearby proteins. Second, electrons captured in the pixels of the CCD's potential wells are able to spill into neighboring pixel wells when a pixel becomes saturated, which produces CCD blooming and loss of image integrity. Thus, masking ensures that pixel saturation and CCD blooming does not occur, guaranteeing better image quality.

Finally, no 2DE gel image analysis software is perfect. DeCyder, with its proprietary watershed, spot-detection algorithm, does not allow the end user to fully understand how the pixel values are used to estimate background contribution. SExtractor and ImageJ are open-source

packages that provide more user control of the data manipulations. SExtractor was designed for the analysis of astronomical data. While the user-controlled parameters allow one to tailor SExtractor to analyze 32-bit, HDR 2DE gel images, there are certain limitations that cannot be addressed without editing sections of the program code. ImageJ does not have advanced image analysis tools for 2DE gels, but it offers a wide variety of tools to analyze raw 32- or 16-bit data, without lossy data compression. The side-by-side comparison of these packages clearly demonstrated the improvements offered by the SI gel imager.

In spite of the aforementioned limitations, the SI gel image is still capable of detecting 10 pg of fluorescently tagged protein. For an average-sized protein of 50 kDa, 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 subfemtomole 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.

We would like to thank Amritha Parthasarathy, Danielle Schlesinger, Raghunandan Avula, Alex Rodriguez, Dagney Cooke, Minh Tam Le, Alex Hurley, and Malachi Blundon for gel electrophoresis; Emily Furbee and Dr. Joseph Ayoob for Drosophila embryo extracts, Lina Song for information on MS; Peter Barnum, Randy Sargent, and Anne Wright for productive image registration discussions; Dr. Marcel Bruchez for discussion about image processing; and Vinitha Ganesan for useful discussions about the manuscript. This work was support by an IDBR grant from the NSF (1063236).

The authors have declared no conflict of interest.

5 References

[1] Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K., Weissman, J. S., *Nature* 2003, *425*, 737–741.

[2] Anderson, N. L., Anderson, N. G., *Mol. Cell. Proteomics* 2002, *1*, 845–867.

[3] Aebersold, R., Mann, M., *Nature* 2003, *422*, 198–207.

[4] Stahl-Zeng, J., Lange, V., Ossola, R., Eckhardt, K., Krek, W., Aebersold, R., Domon, B., *Mol. Cell. Proteomics* 2007, *6*, 1809–1817.

[5] Lange, V., Picotti, P., Domon, B., Aebersold, R., *Mol. Syst. Biol.* 2008, *4*, 222–240.

[6] Zabel, C., Klose, J., *Methods Mol. Biol.* 2009, *519*, 311–338.

[7] Görg, A., Weiss, W., Dunn, M. J., *Proteomics* 2004, *4*, 3665–3685.

[8] Unlu, M., Morgan, M., Minden, J., *Electrophoresis* 1997, *18*, 2071–2077.

[9] Minden, J. S., Dowd, S. R., Meyer, H. E., Stuhler, K., *Electrophoresis* 2009, *30* (Suppl 1), S156–161.

[10] Minden, J., *BioTechniques* 2007, *43*, 739–743.

[11] Marouga, R., David, S., Hawkins, E., *Anal. Bioanal. Chem.* 2005, 669–678.

[12] Srinivasan, V., Lumia, R., *Proceedings of the IEEE International Symposium on Intelligent Control, 1988*. IEEE 1988, Arlington, VA, USA, pp. 196–200.

[13] Erkmen, B. I., *Interplanet. Netw. Prog. Rep* 2011, *42*, 1–23.

[14] Cortizo, E., Yeras, A. M., Lepore, J., Garavaglia, M., *Opt. Lasers Eng.* 2003, *40*, 117–132.

[15] Bertucci, E., Pilu, M., Mirmehdi, M., *Proceedings of the 7th International Conference on Document Analysis and Recognition, 2003*. IEEE 2003, Edinburgh, Scotland, UK, pp. 555–559.

[16] Kwak, Y., MacDonald, L., *Displays* 2000, *21*, 179–194.

[17] Zitova, B., Flusser, J., *Image Vis. Comput.* 2003, 977–1000.

[18] Bradski, G., *Doctor Dobbs J.* 2000, *25*, 120–126.

[19] Hiraoka, Y., Sedat, J. W., Agard, D. A., *Science* 1987, 36–41.

[20] Janesick, J. R., Elliott, T., Collins, S., Marsh, H., Freeman, J., *28th Annual Technical Symposium*. International Society for Optics and Photonics 1984, pp. 2–31.

[21] Sellers, K., Miecznikowski, J., Viswanathan, S., Minden, J., Eddy, W., *Electrophoresis* 2007, *28*, 3324–3332.

[22] Gong, L., Puri, M., Unlü, M., Young, M., Robertson, K., Viswanathan, S., Krishnaswamy, A., Dowd, S. R., Minden, J. S., *Development* 2004, 643–656.

[23] Bertin, E., Arnouts, S., *Astron. Astrophys. Suppl. Ser.* 1996, *117*, 393–404.

[24] Schrimpf, S. P., Weiss, M., Reiter, L., Ahrens, C. H., Jovanovic, M., Malmstrom, J., Brunner, E., Mohanty, S., Lercher, M. J., Hunziker, P. E., Aebersold, R., von Mering, C., Hengartner, M. O., *PLoS Biol.* 2009, e48.

[25] Amersham, *Ettan DIGE User Manual*, Uppsala 2002.

[26] Kang, Y., Techanukul, T., Mantalaris, A., Nagy, J. M., *J. Proteome Res.* 2009, *8*, 1077–1084.

