Application of Two-Dimensional Gel Electrophoresis To Interrogate Alterations in the Proteome of Genetically Modified Crops. 1. Assessing Analytical Validation

ARTICLE *in* JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY · MARCH 2006

Impact Factor: 2.91 · DOI: 10.1021/jf0523566 · Source: PubMed

J. Agric. Food Chem. **2006,** *54,* 2154−2161

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Application of Two-Dimensional Gel Electrophoresis To Interrogate Alterations in the Proteome of Genetically Modified Crops. 1. Assessing Analytical Validation

MARTIN C. RUEBELT,*,§ NANCY K. LEIMGRUBER,§ MARKUS LIPP,§ TRACEY L. REYNOLDS,§ MARGARET A. NEMETH,§ JAMES D. ASTWOOD,§ KARL-HEINZ ENGEL,† AND KLAUS-DIETER JANY‡

Product Safety Center, Monsanto Company, 800 North Lindbergh Boulevard, St. Louis, Missouri 63167; Center for Food and Life Sciences, Chair of General Food Technology, Technical University of Munich, Am Forum 2, 85350 Freising-Weihenstephan, Germany; and Federal Research Centre for Nutrition and Food, Haid-und-Neu-Strasse 9, 76131 Karlsruhe, Germany

Current tools used to assess the safety of food and feed derived from modern biotechnology emphasize the investigation of possible unintended effects caused directly by the expression of transgenes or indirectly by pleiotropy. These tools include extensive multisite and multiyear agronomic evaluations, compositional analyses, animal nutrition, and classical toxicology evaluations. Because analytical technologies are rapidly developing, proteome analysis based on two-dimensional gel electrophoresis (2DE) was investigated as a complementary tool to the existing technologies. A 2DE method was established for the qualitative and quantitative analysis of the seed proteome of *Arabidopsis thaliana* with the following validation parameters examined: (1) source and scope of variation; (2) repeatability; (3) sensitivity; and (4) linearity of the method. The 2DE method resolves proteins with isoelectric points between 4 and 9 and molecular masses (MM) of 6−120 kDa and is sensitive enough to detect protein levels in the low nanogram range. The separation of the proteins was demonstrated to be very reliable with relative position variations of 1.7 and 1.1% for the p*I* and MM directions, respectively. The mean coefficient of variation of 254 matched spot qualities was found to be 24.8% for the gel-to-gel and 26% for the overall variability. A linear relationship ($R^2 >$ 0.9) between protein amount and spot volume was demonstrated over a 100-fold range for the majority of selected proteins. Therefore, this method could be used to interrogate proteome alterations such as a novel protein, fusion protein, or any other change that affects molecular mass, isoelectric point, and/or quantity of a protein.

KEYWORDS: Proteomics; two-dimensional gel electrophoresis; validation; unintended effect; seed proteome; *Arabidopsis*

INTRODUCTION

Profiling techniques such as transcriptomics, proteomics, and metabolomics are discussed as complementary tools for the safety assessment of genetically modified (GM) crops (*1*, *2*). Currently, the core of the safety assessment is the comparison of the GM crop with a traditional counterpart using well-established methods. This comparative approach is known as the concept of substantial equivalence (*3*) and aims to address any effects (intended and unintended) as a result of the genetic modification. The comparison between GM and non-GM crops usually comprises agronomic/phenotypic characteristics, feed performance studies, and crop composition. Agronomic and phenotypic characteristics are very sensitive indices of alterations in metabolism and potential genetic pleiotropy and, hence, robust indicators of equivalence (*4*). Compositional studies are based on an analysis of key macro- and micronutrients, antinutrients, and known toxins and allergens. Feed performance studies with rapidly growing animal species such as broiler chicks are also extremely sensitive bioassays to detect changes in the level of nutrients or antinutrients (*1*, *5*, *6*). Over 50 "biotech" crops have been determined to be as safe and nutritious as their conventional counterparts (*1*). Worldwide, there has been no verifiable unintended toxic or nutritionally harmful effects resulting from the cultivation and consumption of products from GM crops (*7*, *8*) after 10 years of consumption.

Currently, research is being conducted to test the hypothesis that profiling techniques can add useful information to the safety assessment process (*9*−*14*). Profiling technologies allow the

* Corresponding author (e-mail m.ruebelt@gmail.com).
§ Monsanto Company.
† Technical University of Munich.
‡ Federal Research Centre for Nutrition and Food.

simultaneous measurement and comparison of thousands of plant components without prior knowledge of their identity. Three main cell constituent groups are targeted: RNAs, metabolites, and proteins. The combination of these nontargeted approaches is considered to facilitate a more comprehensive approach than the targeted methods and, thus, provide additional opportunities to identify unintended effects. In this context it should be noted that the occurrence of unintended effects is not unique to the application of recombinant DNA techniques. It is an inherent and general phenomenon that occurs frequently in traditional breeding. Due to the common practice of selecting favorable and discarding unwanted lines in the course of breeding programs, documented reports on unintended effects are rare (*2*, *4*).

Proteins are of special interest for the safety assessment because they may be involved in the synthesis of toxins or antinutrients or be a toxin (e.g., phytohemagglutinin), an antinutrient (e.g., protease inhibitor), or an allergen. Proteomics aims to describe the proteome (protein expression) and its changes under the influence of biological perturbations (e.g., disease or mutation) in a comprehensive and quantitative way (*15*). The proteome is the entire complement of the genome and the result of genetic expression, ribosomal synthesis, and proteolytic degradation (*16*). Contrary to the genome, which is constant for an organism, the proteome of an organism is in flux and depends on cell cycle, environmental influences, and tissue/cell type. Although there is no single, fixed proteome, the proteome nevertheless remains a direct product of the genome. Therefore, if a genetic modification affects the genome (e.g., insertional mutation) or gene regulation (pleiotropic effect) of a plant in a way that it changes metabolic pathways or produces a new protein (e.g., inserted gene product, fusion protein, or activation of a silent gene), the proteome could be altered.

The overall objectives of this series of studies (*17*, *18*) were (1) to establish and validate a facile rapid method for partial proteome analysis, (2) to investigate if nontargeted proteomic methods can add new or different useful information to the safety assessment, and (3) to determine if these methods are of general utility in the characterization of strains (e.g., ecotypes, variants, transgenics).

Various methods are available for the qualitative and quantitative comparison of plant proteomes (reviewed, e.g., in refs *19* and *20*). So far, two-dimensional gel electrophoresis (2DE) combined with mass spectrometry (MS) is still the most widely used approach to compare plant proteomes to identify differentially expressed proteins (*21*). Proteins are separated in the first dimension as polypeptides according to their isoelectric point (p*I*) and in the second dimension according to their molecular mass. Specific proteins can be subsequently sequenced and identified by MS, MS/MS, or N-terminal sequencing. 2DE can also reveal alterations in post-translational modification (PTM) of proteins, such as phosphorylation, geranylation, glycosylation, proteolytic cleavage, translational induction, alternative splicing, or any other modification that affects the molecular mass or isoelectric point.

Although 2DE-based proteomics is a powerful tool, it faces some challenges for deployment in a safety assessment: (i) absence of standardized protocols; (ii) lack of validation data or reference data, which could be used to evaluate the power and limitations of a method; and (iii) relative quantitative data versus absolute data achieved with targeted methods such as ELISA. Proteome analysis is also limited by the static examination of dynamic physiological processes.

This paper reports a proteomics method based on 2DE developed and optimized for the qualitative and quantitative analysis of the seed proteome of *Arabidopsis thaliana*. Method validation is a prerequisite to the application of 2DE-based proteomics to GM safety assessment. Therefore, the following validation parameters were examined: (1) source and scope of variation inherent to the developed proteomics approach; (2) precision or repeatability; (3) sensitivity; and (4) linearity of the method.

MATERIALS AND METHODS

Sample Preparation. Mature seeds (30−40 mg) of *A. thaliana*, ecotype Columbia, were ground in screw-cap tubes with a paint shaker-like device (Monsanto Co., St. Louis, MO). The seed samples were precooled at −80 °C for at least 2 h and then shaken for 45 s at 1500 rpm. The sample tubes were placed on dry ice immediately after the grinding process. Extraction buffer [0.7 mL of 7 M urea, 2 M thiourea, 0.75% (w/v) CHAPS, 0.75% (v/v) Triton X-100, 100 mM DTT, 1% (v/v) carrier ampholytes stock, 20% (v/v) 2-propanol, protease inhibitor cocktail Complete (Roche, Mannheim, Germany)] was added to each tube. The carrier ampholytes stock used in the preparation of all solutions was a 2:1:3 mixture of Ampholine 3.5−9.5 (GE Healthcare, Piscataway, NJ), Pharmalyte 5−8 (GE Healthcare), and Bio-Lyte 3−10 (Bio-Rad, Hercules, CA). The samples were mixed on a Nutator (Becton-Dickinson, Sparks, MD) for 1 h at room temperature and subjected to centrifugation at 16000*g* for 5 min at room temperature to remove insoluble material. Finally, the supernatants were removed with a syringe, and after an additional centrifugation step, they were stored in aliquots at −80 °C until analysis.

Protein Quantification. Protein concentration was estimated using the Bio-Rad Protein Assay based on the dye-binding method of Bradford (*22*). Bovine serum albumin (BSA) was used to generate the standard curve. The samples and BSA were diluted in a solution containing 3 M urea and 3 M guanidine−HCl to avoid protein precipitation.

Repeatability. Equal amounts of seeds obtained from six individual *A. thaliana* plants (WT Col-0) were pooled to one seed pool. All *Arabidopsis* plants were grown side-by-side in the same growth chamber. To assess the extract-to-extract and the gel-to-gel variability, three protein extracts were prepared from the seed pool, and for each extract, 2DE was performed in triplicate.

Limit of Detection (LOD) and Linear Range. An *Arabidopsis* seed extract was diluted to seven different protein levels spanning a 100-fold range of total protein load (2.5, 5, 10, 25, 60, 125, and 250 μg of total protein) plus one control dilution (150 μg of total protein). Two external protein standards [BSA and β-lactoglobulin (β-LG)] were spiked into dilutions 1−7 to span a 200-fold range (BSA, 5−1000 ng; β-LG, 4−820 ng). No standard was added to a control dilution. Soybean trypsin inhibitor (STI; p*I* 4.5, MM 21.5 kDa) was added to each dilution at the same protein level (150 ng for each gel) as normalization standard. Three replicates were performed for each dilution.

Two-Dimensional Gel Electrophoresis. Protein extracts were diluted in rehydration buffer containing 7 M urea, 2 M thiourea, 0.75% (w/v) CHAPS, 0.75% (v/v) Triton X-100, 100 mM DTT, 0.3% (v/v) carrier ampholytes stock, 10% (v/v) 2-propanol, 12.5% (v/v) water-saturated isobutanol, protease inhibitor cocktail Complete (Roche), and a trace of bromophenol blue. Thirteen centimeter long immobilized pH gradient gel strips (IPG) with nonlinear pH 3−10 gradients (GE Healthcare) were cut 0.5 cm on both sides to fit in the isoelectrofocusing (IEF) unit. The gel strips were rehydrated for 20−23 h at room temperature using 230 μL of diluted sample (2.5−250 μg of total protein) in customized rehydration trays. The IEF was carried out using a Bio-Rad PROTEAN IEF cell with a controlled cell temperature of 20 °C and with a maximum current of 50 μA/strip. The running conditions were as follows: from 0 to 500 V in 3 h, from 500 to 4000 V in 6 h, and a final phase of 4000 V to a total of 35000 Vh. After IEF was completed, the strips were stored at −80 °C until required for the second dimension. Prior to applying the focused IPG strips to the second dimension, the IPG strips were equilibrated first for 10 min in 50 mM

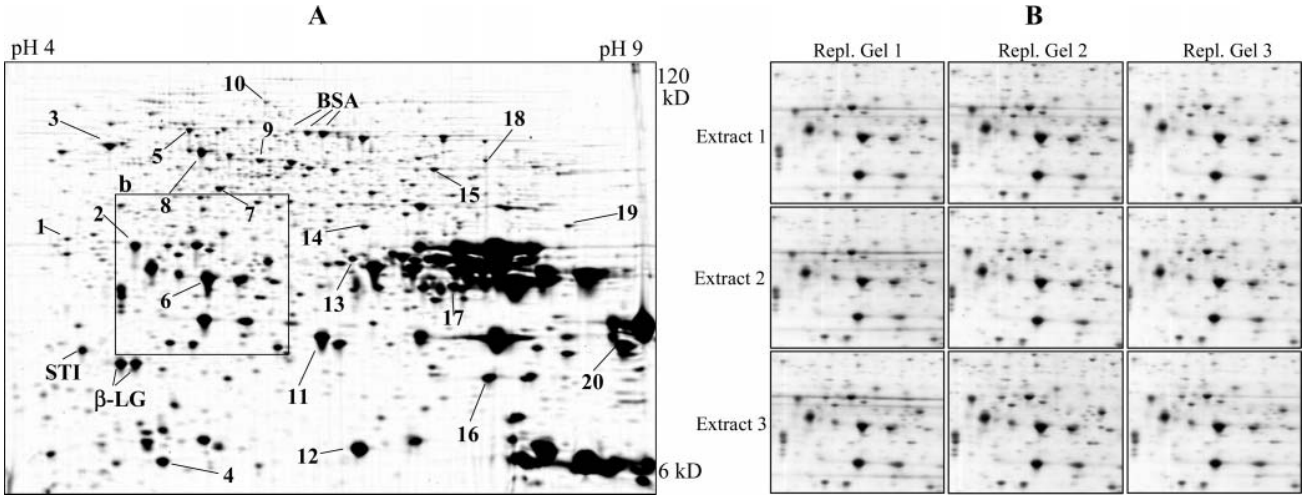

Figure 1. (**A**) Filtered 2DE image of the 125 μg total protein loading. External protein standards and subset member spots are labeled. External protein standards were soybean trypsin inhibitor (STI), bovine serum albumin (BSA), and β-lactoglobulin (β-LG). (**B**) Enlargements of region b of the *A. thaliana* seed 2DE pattern of all nine 2DE gels of the repeatability study. The displayed portions encompass pH 4.7–5.5 and MM 21–42 kDa.

Tris-HCl (pH 8.8), 6 M urea, 30% (w/v) glycerol, 2.3% (w/v) SDS, 1% (w/v) DTT, and bromophenol blue and then for another 10 min in the same solution except DTT was replaced with 4% (w/v) iodoacetamide. The equilibrated IPG strips were then applied to 8−16% Tris-HCl linear gradient Criterion gels and sealed with 1× Tris/glycine/SDS running buffer (Bio-Rad) with 0.9% (w/v) low-melt agarose and a trace of bromophenol blue. The gels were run for the first 15 min at 130 V and then at 180 V until the tracking dye reached the bottom of the gel. Twelve Criterion gels were run at the same time in one Bio-Rad Criterion Dodeca cell for increased reproducibility.

Protein Staining. Colloidal Coomassie Brilliant Blue (CBB) protein staining was performed according to the method of Neuhoff et al. (*23*). Upon completion of electrophoresis, the gels were incubated in fixative (30% methanol, 7% acetic acid) for 1 h. The fixative solution was discarded and replaced with CBB staining solution. The CBB staining solution was prepared fresh by mixing 4 parts of 0.1% (w/v) CBB G-250 in 2% (w/v) phosphoric acid, and 11% w/v ammonium sulfate with 1 part of methanol. The gels were incubated in this solution for 3 days. All incubation steps were performed at room temperature on an orbital rotator. Imaging of the stained proteins was performed at a scan resolution of 36.3 × 36.3 μm using the GS-800 Calibrated Densitometer (Bio-Rad Laboratories).

Image Analysis. The scanned images of the 2DE gels were processed and analyzed with PDQuest 2-DE Gel Analysis software version 7.1 (Bio-Rad Laboratories). The images were cropped and oriented using the image editing controls of the program. All images were processed with the following software settings for spot detection and background subtraction: sensitivity, 40; size scale, 3; min peak, 400; power mean, 3 × 3; floater, 97; speckles filter. These spot detection parameters were chosen as they allowed the detection of the majority of protein spots above the LOD [$LOD = OD_{background} + 3 \times SD_{background}$ (*24*, *25*)] without detecting image noise as spots. Spots detected by the software program were manually verified. False-positive spots (e.g., artifacts and multiple spots in a cluster) were manually removed; false-negative spots (obviously missed spots with OD > LOD) were added to the images. A spot was considered to be reproducibly present/absent when it was present/absent in all three replicate gels of one extraction. To compare spots across gels, a match set was created from the images of the gels in an experiment. A standard gel (Master) was generated out of the image with the greatest number of spots. Spots reproducibly present in a match set member but not present in the image with the most spots were manually added to the standard gel. The automated matching tool of the PDQuest software package (Bio-Rad Laboratories) was used to match spots across the gels. A few landmarks were manually defined to improve the automated matching results. All spots matched by the software program were manually verified. The spots were quantified by 2D Gaussian modeling. Spot quantities of all gels were normalized to remove non-expression-related variations in spot intensity, so the raw quantity of each spot in a gel was divided by the total quantity of all the spots in that gel that have been included in the standard. Data were exported to Excel and from there to JMP for statistical analysis. All statistical analyses were performed with the statistical software package JMP v. 5 (SAS Institute Inc., Cary, NC).

RESULTS AND DISCUSSION

A proteomics method based on 2DE was developed and optimized for the qualitative and quantitative analysis of the seed proteome of *A. thaliana*. The objective was not to further develop and refine proteomics techniques but to set up a robust and reproducible method suitable for routine comparative analysis. Key elements of the developed method include (1) one-step protein extraction, (2) use of commercially available midsize gels (8 × 13 cm) for the first and second dimensions, and (3) colloidal Coomassie Blue staining. Because commercially prepared gels are utilized, the impact from laboratory-specific variation in gel-casting parameters on the result are minimized. The midsize gel format permits high-throughput analysis as the first and second dimensions can be accomplished on the same day and multiple gels (12 per unit) can be run at the same time. Furthermore, midsize gels are more suitable for routine analysis as they are more durable and do not tear as easily as large-format (DALT) gels.

2DE Pattern Repeatability. The impact of the extraction procedure and 2DE on the repeatability (precision) of the spot pattern was investigated by comparing the qualitative (spot presence/absence) and quantitative (spot quantity) variability of the spot patterns among extracts and replicate gels. The 2DE patterns of nine replicate gels were very similar (repeatable) as demonstrated in the displayed enlargements of a randomly picked region of the *Arabidopsis* seed 2DE pattern (**Figure 1B**).

An average of 539 distinct spots was discerned in each of the nine gels. The gel-to-gel standard deviations (SD) (45−129) were much larger than the extract-to-extract standard deviation of 20 (**Table 1**). The large difference between gel-to-gel spot number SD and extract-to-extract spot number SD suggest that the variation in spot number was primarily affected by the 2DE method itself rather than by the sample preparation. Upon visual examination of the 2DE images, focusing problems (streaking), edge effects (pattern distortion), and, in particular,

Table 1. Results of Spot Detection and Matching in the Repeatability Study

	no. of spots					reproducible spots in	
extract	gel 1	gel 2	gel 3	av	SD[a]	replicate gels	all gels
1	434	650	466	517	117	379	350
2	493	578	559	543	45	419	350
3	483	706	482	557	129	403	350
extract-to-extract				539	20	400 ± 20	

[a] Standard deviation.

faint spots on the LOD are the main reasons for the spot number variation. Kamo et al. (*26*) and Gallardo et al. (*27*) were able to resolve 984 and 1272 spots, respectively, from mature *Arabidopsis* seed on DALT 2DE gels. On the basis of DNA sequence data, a total of 25498 proteins is predicted for *Arabidopsis* (*28*). Taking into account post-transcriptional and post-translational regulation such as alternative splicing, phosphorylation, glycosylation, a much higher number of proteins can be expected. Most likely, not all proteins will be expressed and accumulated in seeds and be present in mature seeds. However, this indicates that the 2DE is limited to a subset of the protein population.

Stringent criteria for the presence and absence of a spot are necessary to avoid false positives (artifacts) or negatives. In the present study, a spot was considered to be reproducibly present or absent when the spot was present or absent in all three replicate gels of a sample (extract). To compare the position and the quantitative variation of spots, individual protein spots were matched between all replicate gels. Whereas the number of detected spots ranged from 434 to 706, the number of reproducible spots for the extracts ranged from 379 to 419 (average = 400 ± 20). Three hundred and fifty of these spots were reproducibly present in all nine gels of the three extracts. Spots reproducibly present in one extract were also present in at least one of the replicate gels of the other extract; therefore, the spot patterns of the three extracts did not show any reproducible qualitative difference.

The position of a spot in the gel is another critical qualitative parameter of 2DE and is determined by the isoelectric point and molecular mass of the protein. DNA mutations (point or frameshift mutations) or protein activation changes (post-translational modifications) may lead to changes (shifts) in the isoelectric point or molecular mass of a protein. To detect such protein changes or modifications, the detection of small spot position variations is desirable. To estimate the spot position variation in the optimized 2DE method, the degree of position variation was calculated for the 350 spots matched between all nine gels of the three seed extracts. The average gel-to-gel position standard deviation ($N = 9$) was found to be 1.9 ± 0.4 mm for the *x*-position (isoelectric point) and 0.9 ± 0.2 mm for the *y*-position (molecular mass). By taking into account the dimension of the midsize 2DE gels (110 mm in *x*-position and 80 mm in *y*-position), the relative positional variations are 1.7 and 1.1% for the p*I* and MM directions, respectively. This demonstrates the highly repeatable nature of the 2DE spot position and is consistent with published data (*29*−*34*). It is important to note that the small variation in spot position did not occur independently from surrounding spots. Spot position variation appeared in conjunction with the position variation of neighboring spots as a result of gel distortion. Therefore, an accurate matching of spots and the detection of subtle changes in isoelectric point and molecular mass are possible by considering the spot positions of neighboring spots.

Table 2. Results of the Quantitative Repeatability Study

type of variation	*N*	spots	mean CV (%)	median CV (%)
gel-to-gel	3 × 3	254	24.8 ± 18.5	19.6
overall	9	254	26.2 ± 15.2	22.1

The degree of spot quantity variation inherent to the 2DE process was assessed using only spots matched to all nine gels and with average spot optical densities (OD) above the limit of quantitation [$LOQ = OD_{background} + 10 \times SD_{background}$ (*24*)]. A total of 254 spots (73% of spots present in all nine gels) met these requirements. To accurately compare spot quantities between gels, method-related variations in spot quantity were compensated by normalization. The mean coefficient of variation (CV = SD/mean × 100) is a quantitative index for variation of quantities among matched spots and was computed for gel-to-gel variation and overall variation. The mean CVs of the matched spot quantities for the gel-to-gel and for the overall variability are given in **Table 2**.

Similar quantitative variations were reported in the literature by Mahon et al. (*35*) (mean CV of 32%), Norbeck et al. (*31*) (mean CV of 17%), Blomberg et al. (*30*) (mean CVs of 20−28%), Molloy et al. (*36*) (mean CVs of 18.7−26.4%), and Zhan et al. (*34*) [mean CV of 35.7 ± 20.8% ($n = 3$)]. The small difference of only 1.4% between the gel-to-gel CV and the overall CV indicates that the extract-to-extract variation due to seed grinding and protein extraction has only a minor impact on the quantitative repeatability of the 2DE method. The major contribution to analytical variation results from the 2DE procedure itself. Critical steps of the multistep 2DE method are the application of the sample (in-gel rehydration), the transfer of the proteins from the IPG strip (first dimension) into the SDS-PAGE gel, the two electrophoretic focusing steps, and the staining procedure. The median CVs of spot quantities were lower than their corresponding mean CVs with 19.6% (vs 24.8% mean CV) and 22.1% (vs 26.2% mean CV) for gel-to-gel and overall variability, respectively. This indicates that the mean CV value is affected by the presence of some spots with very poor reproducibility. The CV values ranged from 5.6 to 120% (**Figure 2**). Over 67% of all spots were found to have CVs below 30%, and over 93% of the spots had CVs below 50%. Only 7% (18) of the 254 spots analyzed exhibited a CV of greater than 50%. Closer visual inspection of the 18 spots with CVs over 50% showed that these spots were affected by background, horizontal and vertical streaking, edge effects, and/or neighboring spots and, thus were, inaccurately quantified. Although PDQuest provides a boundary tool to manually define spot contours, which improves the accuracy of spot quantitation (data not shown), this tool is rather time-consuming and, therefore, restricted to a few protein spots of special interest. The extent of a spot CV and the number of replicates determine the sensitivity to detect differences in quantity for this spot. On the basis of the overall averaged variance of all 254 spots of the data set, three replicates allow the assessment of the majority (72%) of spots for a 2-fold change in spot quantity with 90% power and $\alpha = 0.05$.

The 2DE analysis software package PDQuest assigns a spot quality value to each spot ranging from 0 (very bad) to 100 (very good) that is calculated on the basis of shape (Gaussian fit), horizontal streaking, vertical streaking, overlapping, and linear range of scanner. The spot quality and the CV value are good indicators for ambiguous spot quantification. To obtain a high-quality data set and to avoid overestimation of differences,

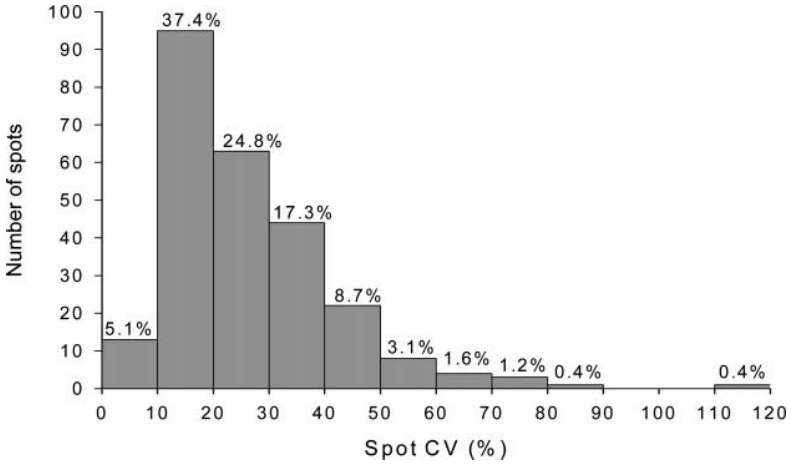

Figure 2. Distribution of the overall spot quantity coefficients of variation for the 254 spots detected in all nine gels and with mean spot quantities above the LOQ.

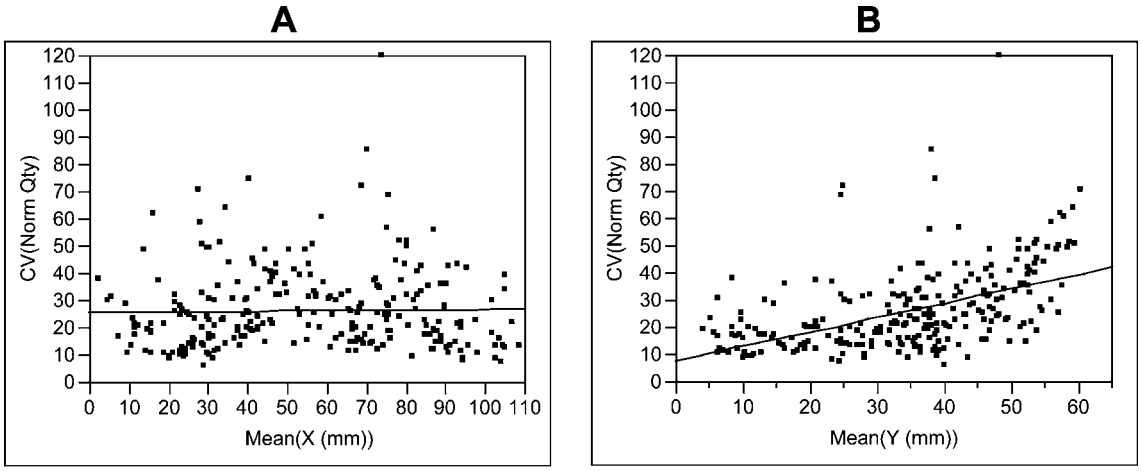

Figure 3. Correlation between spot quantity CV and spot position in the 2DE gel: (**A**) in *x*-position [acidic pH (10 mm = p*I* 4.5) to basic pH (100 mm = p*I* 7.8)]; (**B**) in *y*-position [low MM (5 mm = 6 kDa) to high MM (60 mm = 86 kDa)]. Calculations of CV and *x*- and *y*-position values represent averages from all nine gels (all three extracts). Each graph indicates the best-fit line (or the least-squares regression line).

a spot quality and CV threshold filter should be applied to eliminate inaccurately quantified spots.

To evaluate any potential correlation between variation in spot quantity and spot position, the average spot quantity CV from all spots above the LOQ was related to their spot position in the gel (**Figure 3**). It is apparent that the degree of quantitative variation was evenly distributed in the dimension of isoelectric points (**Figure 3A**). Unlike Norbeck and Blomberg (*31*), who did not see a dependence of spot quantity CV and protein position, a significant ($P < 0.001$) correlation between spot quantity CV and molecular mass was observed (**Figure 3B**). Higher molecular mass proteins showed larger quantitative variation. It is known that the separation of high molecular mass proteins with IPG-2DE is difficult because of size exclusion effects of the IPG gels (*37*−*39*). This may explain why high molecular mass proteins are more susceptible to horizontal streaking, making an accurate quantification of those spots difficult.

Sensitivity and Linearity of Response. The LOD and range of linear response are important for determining the range of proteins this method can evaluate. They are functions of the staining procedure. The published LODs for colloidal Coomassie Brilliant Blue G-250 (CBB G-250) staining vary tremendously from <1 ng (*23*) to up to 100 ng (*40*). Similar discrepancies were found regarding published linear ranges of colloidal CBB. Berggren et al. (*41*) reported a linear range of 8-fold for colloidal CBB stain on 1D SDS-PAGE gels with various standard proteins, including BSA (30−250 ng). Mahon et al. (*35*) determined the linear range by using 2DE and demonstrated a linear range of 20-fold (e.g., BSA, 400 ng to 8 μg). The differences in reported sensitivity and linear ranges may be explained by differences in gel size, gel thickness, and gel type, as well as differences in staining protocols or duration. It is obvious that sensitivity and linear range are method-specific parameters and, therefore, have to be determined individually for every developed method.

To estimate the sensitivity of the developed 2DE method, the LOD was determined for two standard proteins. BSA and β-LG were chosen as external protein standards as they have different molecular masses and are not present in *Arabidopsis* seeds. The two protein standards were spiked into seven *Arabidopsis* seed extract dilutions in different amounts to span a 200-fold range (BSA, 5−1000 ng; β-LG, 4−820 ng). **Figure 1A** shows the locations in the 2DE gel of the external proteins, and **Figure 4** shows the 2DE gel areas containing the two external protein standards for each dilution.

BSA separated into three protein spots (**Figure 4**) with the same MM of 66.7 kDa and p*I* values of 5.4, 5.5, and 5.6. β-LG separated into two protein spots (**Figure 4**) with the same MM of 20.1 kDa and p*I* values of 4.6 and 4.7. The automatic spot

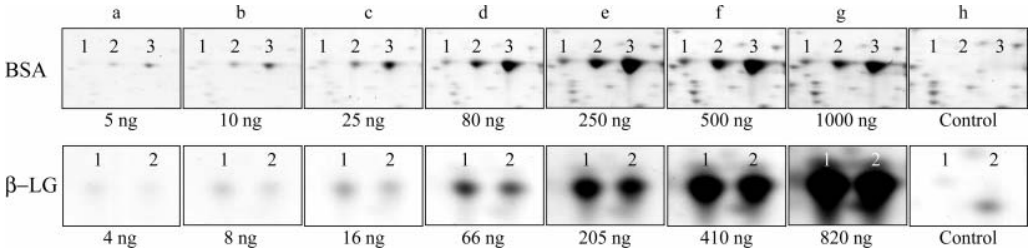

Figure 4. Bovine serum albumin (BSA) and β-lactoglobulin (β-LG) separated by 2DE. BSA appeared as two main spots (2 and 3) and one faint spot (1) with p*I* values of 5.4, 5.5, and 5.6 and a MM of 66.7 kDa. β-LG appeared as two spots (1 and 2) with p*I* values of 4.6 and 4.7 and a MM of 20.1 kDa. The absolute protein load of BSA and β-LG is indicated below the corresponding image. The total *Arabidopsis* seed protein load was for dilution (**a**) 2.5 μg, (**b**) 5 μg, (**c**) 10 μg, (**d**) 25 μg, (**e**) 60 μg, (**f**) 125 μg, (**g**) 250 μg, and (**h**) 150 μg. Images of a dilution series are displayed with the same brightness and contrast settings.

Table 3. Linearity Test between Protein Amount and Staining Intensity for 20 *Arabidopsis* Seed Proteins

spot	p*I*	mol mass (kDa)	tested range (μg of total protein)	linear range (μg of total protein)	R^2	P	R^2 (log/log)	*P* (log/log)
1	4.3	35	10–250	10–250	0.992	0.001	0.994	0.185
2	4.7	34	2.5–250	2.5–250	0.980	0.828		
3	4.6	62	2.5–250	2.5–250	0.985	0.501		
4	4.9	7	2.5–250	2.5–250	0.986	0.549		
5	5.0	69	2.5–250	2.5–60	0.949	0.958		
6	5.1	28	2.5–250	2.5–250	0.986	0.886		
7	5.2	45	2.5–250	2.5–250	0.986	0.988		
8	5.1	58	2.5–250	2.5–250	0.989	0.992		
9	5.4	54	2.5–250	2.5–250	0.946	0.138		
10	5.4	103	2.5–250	2.5–60	0.931	0.227		
11	5.6	22	2.5–250	2.5–250	0.986	0.207		
12	5.7	8	2.5–250	2.5–250	0.966	0.916		
13	5.7	31	2.5–250	2.5–250	0.842	0.543	0.958	0.232
14	5.8	36	2.5–250	2.5–250	0.918	0.115		
15	6.2	50	2.5–250	2.5–60	0.970	0.746		
16	6.4	18	2.5–250	2.5–250	0.921	0.984		
17	6.3	27	2.5–250	2.5–250	0.900	0.763		
18	6.4	54	2.5–250	2.5–60	0.942	0.804		
19	6.9	36	2.5–250	2.5–250	0.856	0.985	0.969	0.597
20	8.6	21	2.5–250	2.5–250	0.934	0.931		

detection tool of PDQuest detected BSA spots 2 and 3 down to 5 ng and spot 1 down to 80 ng BSA. Both spots of β-LG were automatically detected down to 66 ng of β-LG (dilution **d**). At a similar nanogram levels (e.g., 10 ng of BSA vs 8 ng of β-LG) BSA spots 2 and 3 gave a stronger response than spots 1 and 2 of β-LG (**Figure 4**). The discrepancy in staining intensity between BSA and β-LG may result from the Coomassie Blue predominately binding to basic and sulfur-containing amino acids of proteins (*42*). In addition, BSA by weight has more basic amino acids and cysteines than β-LG, and low-MM spots have a higher diffusion coefficient than high-MM spots (*43*), leading to wider spots with weak spot intensities. After image brightness and contrast had been adjusted for each image section individually, all of the spots of BSA and β-LG were clearly distinguishable from the background, even at the lowest tested nanogram levels of 5 and 4 ng for BSA and β-LG, respectively.

Taking into account that the protein amount of 5 ng for BSA is divided by three spots with ratios of approximately 8, 25, and 67% of total spot quantities (quantities of all three spots summed), ~1.3 ng of BSA protein was detectable automatically with PDQuest and as little as 0.4 ng of BSA protein was visually detectable. β-LG separated into two spots with approximate proportions of 43 and 57% of total spot amount; that is, down to 1.7 and 2.3 ng of β-LG were visually detectable by using 2DE combined with colloidal CBB. If proteins present at levels of <1 ng are of interest, specific enrichment techniques can be applied before 2DE.

The linear relationship between protein amount and staining intensity was examined for each of the two spiked standard proteins BSA and β-LG over a 200-fold range (BSA, 5−1000 ng; β-LG, 4−820 ng). The normalized total spot quantity (sum of individual spots for the spiked protein standards) was plotted against protein load to evaluate the linear relationship between spot quantity and protein-loading amount. A linear relationship was considered when the coefficient of determination (R^2) between spot quantity and the protein-loading amount was >0.9 and the lack of fit was insignificant ($P > 0.05$). The linear ranges of BSA and β-LG were found to be 5−250 ng (50-fold; $R^2 = 0.912$) and 8−820 ng (100 fold; $R^2 = 0.981$), respectively. A linear relationship over the same range was also seen when the quantities of the individual spots were plotted against protein amount (data not shown). BSA showed saturation effects at the two highest protein amounts (500 and 1000 ng), and these concentrations were excluded from the calibration plot. Similarly, for β-LG, the 4 ng spot was excluded from the linearity test; its average spot quantity was almost equal to the average spot quantity of the 8 ng spot, and its coefficient of variation was 86%. The demonstrated linear range (up to 100-fold) of loaded protein amount and spot quantity is better than the linear ranges published (up to 20-fold) for colloidal CBB (*35*, *41*). This might be because of the prolonged staining time of the staining protocol, which advances complete protein staining throughout the entire cross section of the gel layer even in areas of high protein density. A large linear range allows for maximal protein quantitation and comparison.

To investigate the linearity for *Arabidopsis* seed proteins, a subset of 20 seed proteins, representing a wide range of different p*I* values (4.3−8.6), MMs (7−103 kD), and abundances was chosen (1−20, **Figure 1A**). The linearity was tested over a 100-fold range. A linear response over the entire detected range was demonstrated for 16 spots (80%) (**Table 3**). Fifteen of these spots were linear over a 100-fold range (2.5−250 μg of total protein) and one over a 25-fold range (10−250 μg of total protein). In three cases (spots 1, 13, and 19), a log transformation of the spot quantities was performed to remove the relationship between the mean and the variance. However, four spots (5, 10, 15, and 18) showed saturation effects and had a limited dynamic range (2.5−60 μg of total protein). All four spots are located in the upper quarter of the 2DE gel (MM > 50 kDa), suggesting a relationship between MM and the saturation effect. A possible explanation may be that high molecular mass proteins have a smaller diffusion coefficient than low molecular mass proteins and tend to be more concentrated in one spot. However, spots with similar molecular masses, such as spots 3 (61.7 kDa), 8 (58.2 kDa), and 9 (53.9 kDa), do not show such saturation effects. Thus, the linear range depends on the protein itself rather than on the MM. Nishihara et al. (*44*) also found differences in linear ranges for different proteins with SYPRO Ruby as stain.

The large dynamic range of proteins in biological samples is a major challenge; it can span up to 8 orders of magnitude (*15*, *38*, *45*). So far, no staining method can accurately quantitate over this range (*46*). To analyze all proteins within the linear range, multiple gels with different protein loads have to be performed. Taking into account that 75% of the tested spots did not show saturation effects over a 100-fold range and that all spots showed a linear response over at least a 25-fold range, it is estimated that the majority of the resolved seed proteins seem to fall in the linear range of the developed method.

In conclusion, a proteomics method, based on two-dimensional gel electrophoresis (2DE), was developed for the qualitative and quantitative analysis of the seed proteome of *A. thaliana* and validated for repeatability, sensitivity, and linearity. The developed 2DE method resolves proteins with isoelectric points between 4 and 9 and molecular masses of 6−120 kDa and is sensitive enough to detect protein levels in the low nanogram range. On average, 500 protein spots could be resolved on a midsize 2DE gel, which represents only a subset of the expected *Arabidopsis* seed proteome. All of the tested seed proteins demonstrated a linear range over 25-fold. Yet, because of the extreme dynamic range of proteins in seeds, it will not be possible to analyze all proteins within their linear range. Hence, a change in protein amount may be undetected (false negative). The spot position is highly repeatable for the isoelectric point and molecular mass dimensions and allowed three BSA isoforms to be separated with p*I* differences of only 0.1 pH unit. Within the spectrum of proteins surveyed, 2DE provides a tool to interrogate proteome alterations such as a novel protein, fusion protein, or any other protein change that affects molecular mass or isoelectric point. Notwithstanding the limitations of the linear range, the 2DE method allows the simultaneous comparison of hundreds of proteins, yet the high gel-to-gel variability demonstrated susceptibility of the 2DE method for random errors. Small variations in this multistep procedure may have major influences on the resulting protein pattern. Hence, the comparison of datasets between different laboratories will be a major challenge, although a standardized method may be developed for sample preparation and electrophoresis. Any standardization procedure should include within laboratory and interlaboratory validation.

LITERATURE CITED

(1) Chassy, B.; Hlywka, J. J.; Kleter, G. A.; Kok, E. J.; Kuiper, H. A.; McGloughlin, M.; Munro, I. C.; Phipps, R. H.; Reid, J. E. Nutritional and safety assessments of foods and feeds nutritionally improved through biotechnology: an executive summary. *Compr. Rev. Food Sci. Food Saf.* **2004**, *3*, 38−104.

(2) Cellini, F.; Chesson, A.; Colquhoun, I.; Constable, A.; Davies, H. V.; Engel, K. H.; Gatehouse, A. M.; Karenlampi, S.; Kok, E. J.; Leguay, J. J.; Lehesranta, S.; Noteborn, H. P.; Pedersen, J.; Smith, M. Unintended effects and their detection in genetically modified crops. *Food Chem. Toxicol.* **2004**, *42*, 1089−1125.

(3) *OECD Safety Evaluation of Foods Derived by Modern Biotechnology: Concepts and Principles*; 1993; http://www.oecd.org/dataoecd/57/53/1946129.pdf.

(4) Cockburn, A. Assuring the safety of genetically modified (GM) foods: the importance of an holistic, integrative approach. *J. Biotechnol.* **2002**, *98*, 79−106.

(5) Konig, A.; Cockburn, A.; Crevel, R. W.; Debruyne, E.; Grafstroem, R.; Hammerling, U.; Kimber, I.; Knudsen, I.; Kuiper, H. A.; Peijnenburg, A. A.; Penninks, A. H.; Poulsen, M.; Schauzu, M.; Wal, J. M. Assessment of the safety of foods derived from genetically modified (GM) crops. *Food Chem. Toxicol.* **2004**, *42*, 1047−1088.

(6) European Food Safety Authority Guidance document of the Scientific Panel on Genetically Modified Organisms for the risk assessment of genetically modified plants and derived food and feed. *EFSA J.* **2004**, *99*, 1−94.

(7) *GM Science Review Panel—UK GM Science Review Panel: Second Report*; an open review of the science relevant to GM crops and food based on interests and concerns of the public; Jan 22, 2004; http://www.gmsciencedebate.org.uk/report/.

(8) Society of Toxicology. The safety of genetically modified foods produced through biotechnology. *Toxicol. Sci.* **2003**, *71*, 2−8.

(9) Kuiper, H. A.; Kok, E. J.; Engel, K. H. Exploitation of molecular profiling techniques for GM food safety assessment. *Curr. Opin. Biotechnol.* **2003**, *14*, 238−243.

(10) Kok, E. J.; Kuiper, H. A. Comparative safety assessment for biotech crops. *Trends Biotechnol.* **2003**, *21*, 439−444.

(11) FAO/WHO. *Safety Aspects of Genetically Modified Foods of Plant Origin. Joint FAO/WHO Consultation on Foods Derived from Biotechnology*, Geneva, Switzerland, May 29−June 2, 2000; Food and Agriculture Organization of the United Nations and World Health Organization: Rome, Italy, 2000.

(12) Kuiper, H. A.; Kleter, G. A.; Noteborn, H. P. J. M.; Kok, E. J. Assessment of the food safety issues related to genetically modified foods. *Plant J.* **2001**, *27*, 503−528.

(13) *EC Scientific Steering Committee Guidance Document for the Risk Assessment of Genetically Modified Plants and Derived Food and Feed* (March 6−7, 2003); European Commission: Brussels, Belgium, 2003; http://europa.eu.int/comm/food/fs/sc/ssc/out327_en.pdf.

(14) *National Research Council Safety of Genetically Engineered Foods: Approaches to Assessing Unintended Health Effects*; National Academies Press: Washington, DC, 2004; http://books.nap.edu/catalog/10977.html.

(15) Anderson, N. L.; Anderson, N. G. Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* **1998**, *19*, 1853−1861.

(16) Wilkins, M. R.; Sanchez, J. C.; Gooley, A. A.; Appel, R. D.; Humphery-Smith, I.; Hochstrasser, D. F.; Williams, K. L. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol. Genet. Eng. Rev.* **1996**, *13*, 19−50.

(17) Ruebelt, M. C.; Lipp, M.; Reynolds, T. L.; Astwood, J. D.; Engel, K.-H.; Jany, K.-D. Application of two-dimensional gel electrophoresis to interrogate alterations in the proteome of genetically modified crops. 2. Assessing natural variability. *J. Agric. Food Chem.* **2006**, *54*, 2162−2168.

(18) Ruebelt, M. C.; Lipp, M.; Reynolds, T. L.; Schmuke, J. J.; Astwood, J. D.; DellaPenna, D.; Engel, K.-H.; Jany, K.-D.

Application of two-dimensional gel electrophoresis to interrogate alterations in the proteome of genetically modified crops. 3. Assessing unintended effects. *J. Agric. Food Chem.* **2006**, *54*, 2169−2177.
(19) van Wijk, K. J. Challenges and prospects of plant proteomics. *Plant Physiol.* **2001**, *126*, 501−508.
(20) Canovas, F. M.; Dumas-Gaudot, E.; Recorbet, G.; Jorrin, J.; Mock, H. P.; Rossignol, M. Plant proteome analysis. *Proteomics* **2004**, *4*, 285−298.
(21) Rabilloud, T. Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains. *Proteomics* **2002**, *2*, 3−10.
(22) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein−dye binding. *Anal. Biochem.* **1976**, *72*, 248−254.
(23) Neuhoff, V.; Arold, N.; Taube, D.; Ehrhardt, W. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* **1988**, *9*, 255−262.
(24) Miller, J. N.; Miller, J. C. *Statistics and Chemometrics for Analytical Chemistry*, 4th ed.; Pearson Education: London, U.K., 2000; p 288.
(25) Fajgelj, A.; Ambrus, A. *Principles and Practices of Method Validation*. The Royal Society of Chemistry: Cambridge, U.K., 2000; Vol. 256.
(26) Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A. Separation and characterization of *Arabidopsis thaliana* proteins by two-dimensional gel electrophoresis. *Electrophoresis* **1995**, *16*, 423−430.
(27) Gallardo, K.; Job, C.; Groot, S. P.; Puype, M.; Demol, H.; Vandekerckhove, J.; Job, D. Proteomic analysis of *Arabidopsis* seed germination and priming. *Plant Physiol.* **2001**, *126*, 835−848.
(28) Bevan, M.; Mayer, K.; White, O.; Eisen, J. A.; Preuss, D.; Bureau, T.; Salzberg, S. L.; Mewes, H. W. Sequence and analysis of the *Arabidopsis* genome. *Curr. Opin. Plant Biol.* **2001**, *4*, 105−110.
(29) Corbett, J. M.; Dunn, M. J.; Posch, A.; Gorg, A. Positional reproducibility of protein spots in two-dimensional polyacrylamide gel electrophoresis using immobilised pH gradient isoelectric focusing in the first dimension: an interlaboratory comparison. *Electrophoresis* **1994**, *15*, 1205−1211.
(30) Blomberg, A.; Blomberg, L.; Norbeck, J.; Fey, S. J.; Larsen, P. M.; Larsen, M.; Roepstorff, P.; Degand, H.; Boutry, M.; Posch, A.; Gorg, A. Interlaboratory reproducibility of yeast protein patterns analyzed by immobilized pH gradient two-dimensional gel electrophoresis. *Electrophoresis* **1995**, *16*, 1935−1945.
(31) Norbeck, J.; Blomberg, A. Two-dimensional electrophoretic separation of yeast proteins using a non-linear wide range (pH 3−10) immobilized pH gradient in the first dimension; reproducibility and evidence for isoelectric focusing of alkaline (pI > 7) proteins. *Yeast* **1997**, *13*, 1519−1534.
(32) Li, C.; Chen, Z.; Xiao, Z.; Wu, X.; Zhan, X.; Zhang, X.; Li, M.; Li, J.; Feng, X.; Liang, S.; Chen, P.; Xie, J. Y. Comparative proteomics analysis of human lung squamous carcinoma. *Biochem. Biophys. Res. Commun.* **2003**, *309*, 253−260.
(33) Zhan, X.; Desiderio, D. M. A reference map of a human pituitary adenoma proteome. *Proteomics* **2003**, *3*, 699−713.
(34) Zhan, X.; Desiderio, D. M. Differences in the spatial and quantitative reproducibility between two second-dimensional gel electrophoresis systems. *Electrophoresis* **2003**, *24*, 1834−1846.
(35) Mahon, P.; Dupree, P. Quantitative and reproducible two-dimensional gel analysis using Phoretix 2D Full. *Electrophoresis* **2001**, *22*, 2075−2085.
(36) Molloy, M. P.; Brzezinski, E. E.; Hang, J.; McDowell, M. T.; VanBogelen, R. A. Overcoming technical variation and biological variation in quantitative proteomics. *Proteomics* **2003**, *3*, 1912−1919.
(37) Oh-Ishi, M.; Satoh, M.; Maeda, T. Preparative two-dimensional gel electrophoresis with agarose gels in the first dimension for high molecular mass proteins. *Electrophoresis* **2000**, *21*, 1653−1669.
(38) Corthals, G. L.; Wasinger, V. C.; Hochstrasser, D. F.; Sanchez, J. C. The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* **2000**, *21*, 1104−1115.
(39) Ong, S. E.; Pandey, A. An evaluation of the use of two-dimensional gel electrophoresis in proteomics. *Biomol. Eng.* **2001**, *18*, 195−205.
(40) Rabilloud, T. Detecting proteins separated by 2-D gel electrophoresis. *Anal. Chem.* **2000**, *72*, 48A−55A.
(41) Berggren, K.; Chernokalskaya, E.; Steinberg, T. H.; Kemper, C.; Lopez, M. F.; Diwu, Z.; Haugland, R. P.; Patton, W. F. Background-free, high sensitivity staining of proteins in one- and two-dimensional sodium dodecyl sulfate-polyacrylamide gels using a luminescent ruthenium complex. *Electrophoresis* **2000**, *21*, 2509−2521.
(42) Sapan, C. V.; Lundblad, R. L.; Price, N. C. Colorimetric protein assay techniques. *Biotechnol. Appl. Biochem.* **1999**, *29* (Part 2), 99−108.
(43) Appel, R. D.; Palagi, P. M.; Walther, D.; Vargas, J. R.; Sanchez, J. C.; Ravier, F.; Pasquali, C.; Hochstrasser, D. F. Melanie II—a third-generation software package for analysis of two-dimensional electrophoresis images: I. Features and user interface. *Electrophoresis* **1997**, *18*, 2724−2734.
(44) Nishihara, J. C.; Champion, K. M. Quantitative evaluation of proteins in one- and two-dimensional polyacrylamide gels using a fluorescent stain. *Electrophoresis* **2002**, *23*, 2203−2215.
(45) Pradet-Balade, B.; Boulme, F.; Beug, H.; Mullner, E. W.; Garcia-Sanz, J. A. Translation control: bridging the gap between genomics and proteomics? *Trends Biochem. Sci.* **2001**, *26*, 225−229.
(46) Wirth, P. J.; Romano, A. Staining methods in gel electrophoresis, including the use of multiple detection methods. *J. Chromatogr. A* **1995**, *698*, 123−143.

Received for review September 26, 2005. Revised manuscript received January 19, 2006. Accepted January 20, 2006.

JF0523566