POLISH JOURNAL OF FOOD AND NUTRITION SCIENCES

http://journal.pan.olsztyn.pl
e-mail: pjfns@pan.olsztyn.pl

Pol. J. Food Nutr. Sci.
2009, Vol. 59, No. 2, pp. 141-144

REPRODUCIBILITY OF TWO-DIMENSIONAL ELECTROPHORESIS GEL IMAGES OF PEA (*PISUM SATIVUM* L.) SEED PROTEINS EVALUATED USING SCATTER PLOTS – A SHORT REPORT

Jerzy Dziuba, Piotr Minkiewicz, Dorota Nałęcz, Marta Dziuba, Iwona Szerszunowicz

Chair of Food Biochemistry, University of Warmia and Mazury in Olsztyn, Poland

Key words: pea seed proteins, 2D-PAGE, reproducibility, image analysis, scatter plot

The aim of this study was to evaluate plots showing a linear correlation between the parameters characterising protein content in the spots obtained using two-dimensional electrophoresis (scatter plots). The correlation coefficient of normalized spot volume may be used to determine regions with increased reproducibility. In our experiment, conducted using pea seed proteins, the correlation coefficient reached 0.980 in such regions, compared with 0.917 noted for the entire gel surface. The correlation coefficient calculated based on scatter plots may serve as a tool for the preliminary selection of regions characteristic of a given proteome.

INTRODUCTION

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) is a basic separation technique applied in proteomic studies including application for food and nutrition science [Hamdan & Righetti, 2003; Kvasnièka, 2003; Görg *et al.*, 2004; Manso et *al.*, 2005; Bendixen, 2005; Skylas *et al.*, 2005; Wang *et al.*, 2006a; Fenselau, 2007]. When the application of 2D-PAGE is followed by mass spectrometry, spots occurring on several gels are usually analysed. The reproducibility of gel images is a crucial factor in such experiments. Gel images may also be classified and discriminated using sophisticated statistical and chemometrical algorithms [Pietrogrande *et al.*, 2006; Marengo *et al.*, 2007; Jacobsen *et al.*, 2007]. Reproducibility estimation and the identification of factors causing gel-to-gel variation is a significant area of interest [Choe & Lee, 2003; Gustaffson *et al.*, 2004; Almeida *et al.*, 2005; Grove *et al.*, 2006; Horgan, 2007; Valcu & Valcu, 2007; Schröder *et al.*, 2008]. The routine use of two-dimensional electrophoresis requires a rapid test to estimate image reproducibility. Such a test should be as simple as possible. A scatter plot, *i.e.* a plot highlighting a linear correlation between such parameters as spot intensity, spot area, spot volume, normalized intensity or normalized volume for matched spots found on two gels [Image Master Manual, 2005] appears to be a simple and fast tool to compare the properties of images. The aim of the present experiment was to evaluate the so-called "scatter plots", the option provided by software for gel image processing as a tool for image reproducibility estimation and to highlight some consequences of its use.

MATERIALS AND METHODS

Seeds of pea (*Pisum sativum* L.) cv. Ramrod, harvested in 2005, were supplied by the Plant Breeding Station in Piaski-Szelejewo (Poland). Seeds were dehulled and milled into flour using a WZ-1 laboratory mill (Spomasz, Żnin, Poland) and a Fack S-2601 tecator (Höganas, Sweden) with a grid diameter <150 μm. Flour was stored in hermetically closed plastic containers at a temperature of around 5°C. Only electrophoresis-purity reagents were used.

Protein extraction was performed by the method proposed by Wang *et al.* [2003; 2006b]. Protein concentration was measured according to Bradford [1976] using the Bio-Rad Protein Assay (Bio-Rad, cat. No. 500-0006) and bovine serum albumin (Bio-Rad, cat. No. 500-0007) as a standard protein.

2D-PAGE was carried out according to Görg [2004]. Samples of the resultant solutions (350 μL, containing *ca.* 300 μg of protein) were loaded onto IPG Dry Strips; 18 cm; pH 3-10 linear gradient (Bio-Rad, cat. No. 163-2032). Isoelectric focusing (IEF) was carried out using Ettan™ IPGphor™ (GE Healthcare). The strips were then transferred to 12.5% polyacrylamide gels (215 × 276 × 1 mm) and covered with a 0.5% agarose solution in the Ettan™ DALTsix (GE Healthcare) apparatus. 2-D SDS-PAGE Standards (Bio-Rad, cat. No. 161-0310) were used.

SDS-PAGE gels were stabilized and stained using Roti®-Blue solution (Colloidal Coomassie G-250, Roth, cat. No. A152.1, Germany). After staining the gels were washed to destain the background. Gels were finally rinsed with deionized water and scanned. The analysis was performed in five replications.

Author's address for correspondence: Prof. Dr. Jerzy Dziuba, Chair of Food Biochemistry, University of Warmia and Mazury in Olsztyn, Plac Cieszyński 1, 10-726 Olsztyn-Kortowo, Poland; tel.: (48 89) 523 35 85; e-mail: jerzy.dziuba@uwm.edu.pl

Gel images were processed using the Image Master™ 2D Platinum 6.0 program [Image Master Manual, 2005]. Spots were detected based on the following parameters selected arbitrarily: minimum area 20; smooth 10; saliency 150. Manual processing was restricted to remove artifacts (*e.g.* bubbles, dust, damages of gel surface) recognized as spots. Apparent pI was calculated using markers identifying the location of both ends of an IPG strip (Figure 1). Apparent molecular weights were calculated on the basis of standards labeled as shown in Figure 1. pI markers and molecular weight standards were used as anchors for matching gel images. Scatter plots of % volume were constructed for matched spots with the apparent pI and MW ranges presented in Table 1. pI markers and MW standards were not taken into account during calculations regarding the entire gel surface.

The statistical significance of differences between parameters describing scatter plots obtained for selected pI and MW ranges and for entire gel was estimated using Student's t-test.

RESULTS AND DISCUSSION

An example of a 2D-PAGE image of pea proteins is presented in Figure 1. A characteristic feature of the image is the presence of abundant proteins in the region corresponding to pI in a neutral and slightly acidic pH range and to the apparent molecular weight range between *ca.* 30 and *ca.* 80 kDa. The value of 80 kDa exceeds the highest molecular weight of standard proteins (Figure 1), thus the region with cutoff 70 kDa was selected for calculations. Most spots in this region are arranged into the so-called "spot trains" containing spots of proteins with similar molecular weights and different pI values. The occurrence of "spot trains" is attributed to the post-translational modifications of proteins and considered to be a characteristic element enabling to identify 2D electrophoregrams [Pietrogrande *et al.*, 2006]. The pattern of spots is similar to that observed in electrophoregrams of soybean seed proteins obtained by Zarkadas *et al.* [2007a, b].

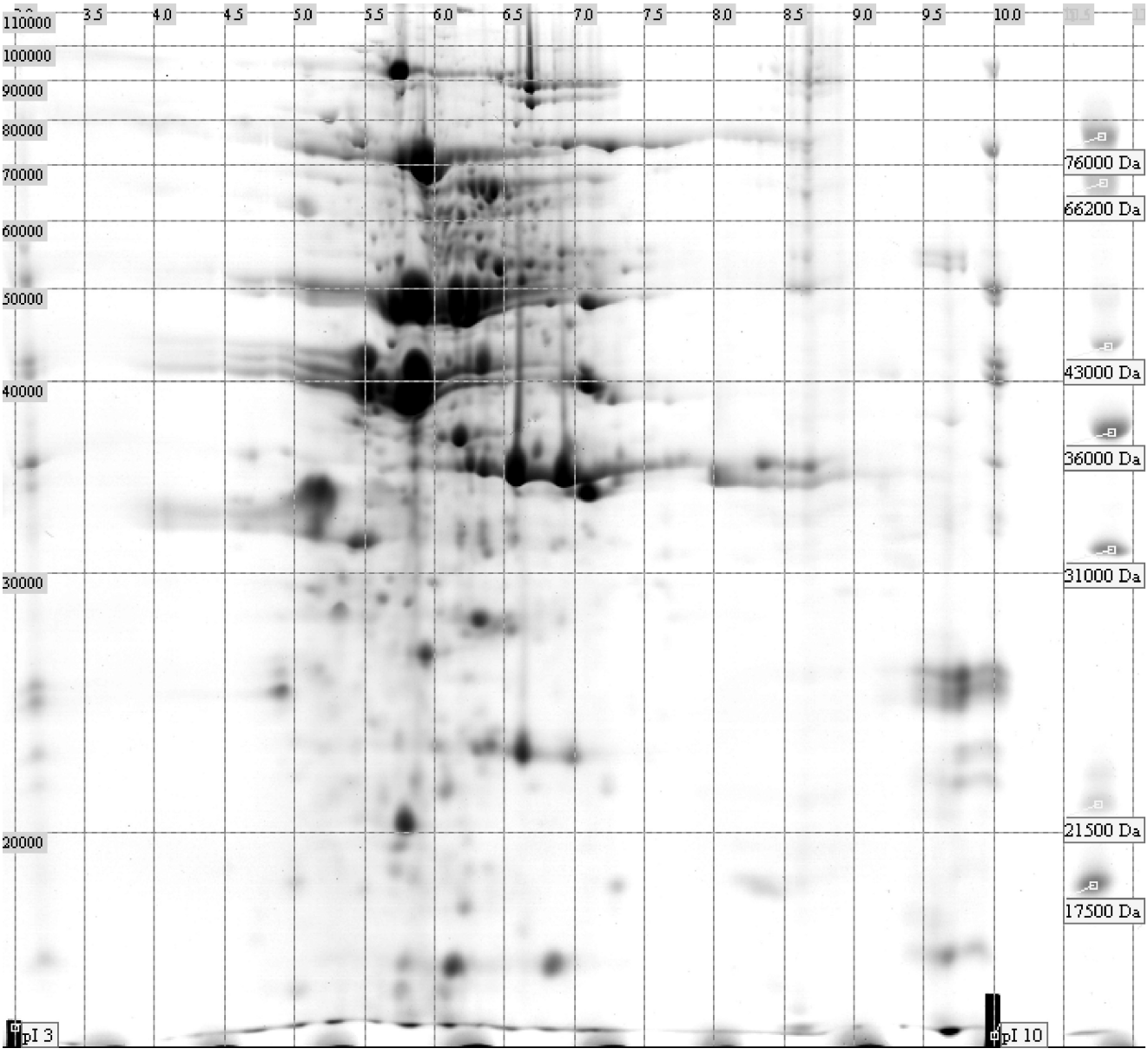

FIGURE 1. Example of an image of a 2D-PAGE electrophoregram of pea seed proteins.

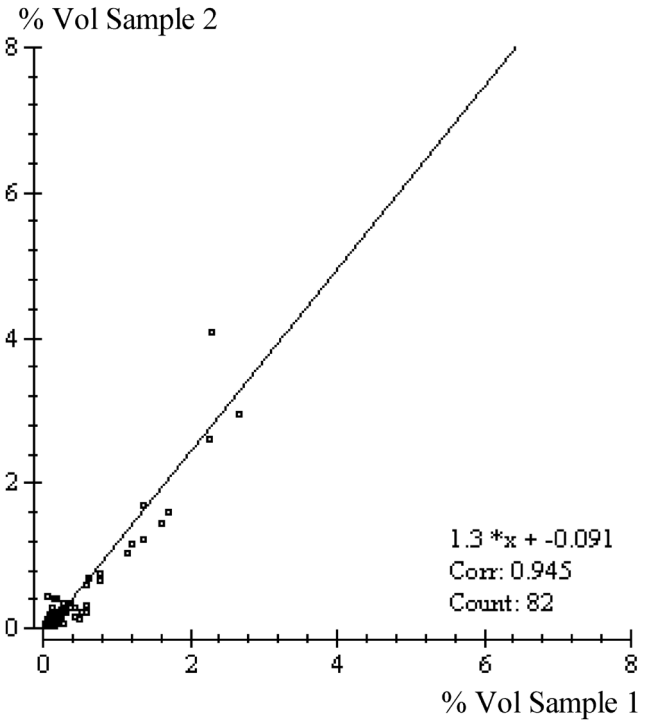

FIGURE 2. Example of a scatter plot for two images within the apparent pI range of 5-7 and the apparent molecular weight range of 40-70 kDa.

An example of a scatter plot is presented in Figure 2. An ideal linear correlation gives correlation coefficient = 1, slope = 1 and offset = 0. The deviation of individual data points from the straight line may be caused by two reasons. The first one is spot overlapping [Pietrogrande *et al.*, 2006]. Overlapped spots may be recognized as one spot in one gel and as two or more in another. The other reason may be the local geometrical distortion of gel [Aittokalio *et al.*, 2005] leading to the incorrect matching of two spots differing in volume. The parameters characterising scatter plots are presented in Table 1. The calculations provided two data sets, with a slope higher than 1 and lower than 1. Only the first one is presented in Table 1. Values from the second set are the inverses of those from the first one. Offset values and correlation coefficients associated with the slopes higher than 1 are presented in Table 1. The program used in the study offers scatter plot calculation for the following parameters: spot intensity, spot area, spot volume, normalized intensity, normalized volume. The last one has given the highest correlation coefficient measured for the entire gel surface, as compared with the other parameters (data not shown). Normalized volume (% volume) is intuitionally the best approximation of the relative protein content of a complex mixture. Among the three measures of correlation between two gels, the correlation coefficient seems to show the most interesting properties. The narrowing of the area covered for spot matching leads to statistically significant changes of this measure (Table 1). It implies that gel images contain areas with reproducibility higher than that of the entire gel image. In our samples, the slope and offset did not point to the fact that some fragments of a gel image show higher reproducibility than others. There was no statistical significance between the values obtained for the entire gel surface and for narrower areas, with one exception (Table 1). In contrast to the correlation coefficient, the slope and offset of normalized spot volume in a narrow area are affected by the volume of all spots in gel. Exceptionally high slopes for data based on matching in the pI range of 6.5-7.0 and in the molecular weight range of 65-70 kDa reflect low robustness of data due to a too low number of spots matched.

The results of this study confirmed that "spot trains" containing intense spots [Pietrogrande *et al.*, 2006] can be recommended for the identification and classification of 2D electrophoresis images.

The correlation coefficients calculated for scatter plots could be a useful tool for the preliminary selection of marker regions indicating changes in proteomes. Regions of increased reproducibility could be selected for further, more sophisticated, analyses.

CONCLUSION

Images obtained by two-dimensional electrophoresis contain regions with reproducibility higher than the overall reproducibility of gels. Linear correlation coefficients calculated for scatter plots may serve as a quantitative measure for the selection of such regions.

TABLE 1. Parameters characterizing scatter plots of normalized spot volume over the entire gel surface and in a selected area.

Apparent pI range	Apparent MW range (kDa)	Number of matches[a] Mean±SD[b]	Correlation coefficient Mean±SD[b]	Slope Mean±SD[b]	Offset Mean±SD[b]
Entire gel surface	Entire gel surface	341±42	0.917±0.019	1.10±0.06	-0.028±0.014
5.0-7.0	30-70	99±15	0.948±0.028[c]	1.15±0.12	-0.027±0.044
5.0-7.0	40-70	78±8	0.953±0.023[c]	1.15±0.12	-0.050±0.026
6.0-7.0	40-70	57±6	0.963±0.027[d]	1.18±0.12	-0.044±0.067
6.0-7.0	50-70	43±5	0.980±0.005[e]	1.17±0.08	-0.018±0.050
6.0-7.0	60-70	26±3	0.968±0.011[e]	1.17±0.08	-0.013±0.060
6.0-7.0	65-70	19±2	0.974±0.011[e]	1.17±0.08	-0.016±0.058
6.5-7.0	65-70	10±1	0.914±0.024	1.40±0.28[c]	-0.031±0.022

[a]Mean number of matches between two individual gels. [b]n = 6; it is possible to obtain six independent scatter plots on the basis of four gel images. [c]Difference between this value and the value obtained for the entire gel surface is significant at $p < 0.05$. [d]Difference between this value and the value obtained for the entire gel surface is significant at $p < 0.01$. [e]Difference between this value and the value obtained for the entire gel surface is significant at $p < 0.001$.

ACKNOWLEDGEMENTS

This work was supported from the own funds of UWM by research projects No. 0712-908 and 0712-203.

REFERENCES

1. Aittokalio T., Salumi J., Nyman T.A., Nevalainen O.S., Geometrical distortions in two-dimensional gels: applicable correction methods. J. Chromatogr. B, 2005, 815, 25–37.
2. Almeida J.S., Stanislaus R., Krug E., Arthur J.M., Normalization and residual variation in two-dimensional gel electrophoresis for quantitative differential proteomics. Proteomics, 2005, 5, 1242––1249.
3. Bendixen E., The use of proteomics in meat science. Meat Sci., 2005, 71, 138–149.
4. Bradford M.M., A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem., 1976, 72, 248–254.
5. Choe L.H., Lee K.H., Quantitative and qualitative measure of intralaboratory two-dimensional protein gel reproducibility and the effects of sample preparation, sample load and image analysis. Electrophoresis, 2003, 24, 3500–3507.
6. Fenselau C., A review of quantitative methods for proteomic studies. J. Chromatogr. B, 2007, 855, 14–20.
7. Görg A., 2-D electrophoresis. Principles and methods. Handbooks from GE Healthcare. 2004, GE Healthcare, Uppsala, Sweden.
8. Görg A., Weiss W., Dunn M.J., Current two-dimensional electrophoresis technology for proteomics. Proteomics, 2004, 4, 3665–3685.
9. Grove H., Hollung K., Uhlen A.K., Martens H., Fægerstad E.M., Challenges related to analysis of protein spot volumes from two-dimensional gel electrophoresis as revealed by replicate gels. J. Proteome Res., 2006, 5, 3399–3410.
10. Gustafsson J.S., Ceasar R., Glasbey C.A., Blomberg A., Rudemo M., Statistical exploration of variation in quantitative two-dimensional electrophoresis data. Proteomics, 2004, 4, 3791–3799.
11. Hamdan M., Righetti P.G., Assessment of protein expression by means of 2-D gel electrophoresis with and without mass spectrometry. Mass Spectrom. Rev., 2003, 22, 272–284.
12. Horgan G.W., Sample size and replication in 2D gel electrophoresis studies. J. Proteome Res., 2007, 6, 2884–2887.
13. Image Master 2D Platinum 6.0 User Manual (2005) GE Healthcare, Uppsala, Sweden (pdf file available directly via Google search).
14. Jacobsen S., Grove H., Jensen K.N., Helle A. Sørensen H.A., Jessen F., Hollung K., Uhlen A.K., Jørgensen B.M., Færgestad E.M., Søndergaard I., Proteomics and 2-DE. Multivariate analysis of 2-DE protein patterns – practical approaches. Electrophoresis, 2007, 28, 1289–1299.
15. Kvasnička F., Review – Proteomics: general strategies and application to nutritionally relevant proteins. J. Chromatogr. B, 2003, 787, 77–91.
16. Manso M.A., Léonil J., Jan G., Gagnaire V., Application of proteomics to characterization of milk and dairy products. Int. Dairy J., 2005, 15, 845–855.
17. Marengo E., Robotti E., Bobba M., Multivariate statistical tools for the evaluation of proteomic 2D-maps: recent achievements and applications. Curr. Proteom., 2007, 4, 53–66.
18. Pietrogrande M.C., Marchetti N., Dondi F., Righetti P.G., Decoding 2D-PAGE complex maps: relevance to proteomics. J. Chromatogr. B, 2006, 833, 51–62.
19. Schröder S., Zhang H., Yeung E.S., Jänsch L., Zabel C., Wätzig H., Quantitative gel electrophoresis: sources of variation. J. Proteome Res., 2008, 7, 1226–1234.
20. Skylas D.J., Van Dyk D., Wrigley C. W., Proteomics of wheat grain. J. Cereal Sci., 2005, 41, 165–179.
21. Valcu C.-M., Valcu M., Reproducibility of two-dimensional gel electrophoresis at different replication levels. J. Proteome Res., 2007, 6, 4677–4683.
22. Wang J.J., Li D.F., Dangott L.J., Wu G.Y., Proteomics and its role in nutrition research. J. Nutr., 2006a, 136, 1759–1762.
23. Wang W., Nagnani R., Scali M., Cresti M., A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. Electrophoresis, 2006b, 27, 2782–2786.
24. Wang W., Scali M., Vignani R., Spadafora A., Sensi E., Mazzuca S., Cresti M., Protein extraction for two-dimensional electrophoresis from olive leaf, a plant tissue containing high levels of interfering compounds. Electrophoresis, 2003, 24, 2369–2375.
25. Zarkadas C.G., Gagnon C., Poysa V., Khanizadeh S., Cober E.R., Chang V., Gleddie S., Protein quality and identification of the storage protein subunits of tofu and mill soybean genotypes, using amino acid analysis, one- and two-dimensional gel electrophoresis and tandem mass spectrometry. Food Res. Int., 2007a, 40, 111–128.
26. Zarkadas C.G., Gagnon C., Gleddie S., Khanizadeh S., Cober E.R., Guillemette R.J.D., Assessment of the protein quality of fourteen soybean [*Glycine max* (L.) Merr.] cultivars using amino acid analysis and two-dimensional electrophoresis. Food Res. Int., 2007b, 40, 129–146.

Received October 2008. Revision received and accepted March 2009.