Published in final edited form as:

Electrophoresis. 2009 March ; 30(5): 848–851. doi:10.1002/elps.200800583.

GelBandFitter – A computer program for analysis of closely spaced electrophoretic and immunoblotted bands

Mihail I. Mitov¹, Marion L. Greaser², and Kenneth S. Campbell¹

¹Department of Physiology, University of Kentucky, Lexington, KY, USA

²Muscle Biology Laboratory, University of Wisconsin-Madison, Madison, WI, USA

Abstract

GelBandFitter is a computer program that uses non-linear regression techniques to fit mathematical functions to densitometry profiles of protein gels. This allows for improved quantification of gels with partially overlapping and potentially asymmetric protein bands. The program can also be used to analyze immunoblots with closely spaced bands. GelBandFitter was developed in Matlab and the source code and/or a Windows executable file can be downloaded at no cost to academic users from <http://www.gelbandfitter.org>.

Keywords

Computer program; Gel analysis; Immunoblot analysis; Myosin

Researchers often wish to calculate the relative amounts of two proteins in an experimental sample. This can be accomplished by imaging a stained gel and marking separate regions of interest around the bands corresponding to the two chosen proteins. The relative amount of each protein can then be determined from the integrals of the densitometry profiles for the two regions.

When the proteins are fully resolved on the gel this procedure is straightforward because the densitometry profile drops to baseline between the two bands and the regions of interest can be defined without significant ambiguity. The approach is more problematic when, as is often the case in real experiments, the two bands partially overlap. When this occurs, the two regions of interest cannot be defined without at least some ambiguity concerning (i) the best way of splitting the convolved profile and (ii) whether the demarcation will bias the experimental results.

GelBandFitter is a computer program that was developed using the Matlab language (The Mathworks, Natick, MA) to provide a better way of analyzing overlapping bands. The program has a graphical user interface and works by using non-linear regression techniques to fit overlapping mathematical functions to the entire densitometry profile. This eliminates the need to identify distinct regions of interest for each band and thereby reduces experimental error. The software can be downloaded at no cost to academic users from

<http://www.gelbandfitter.org> as (i) a collection of source files or (ii) a single executable file for Windows-based computers.

Users who download the source files will need to have their own version of Matlab to run the software (many universities have site licenses for the application) but will be able to modify the program for their own specific needs if they desire. Scientists who download the executable version do not need their own copy of Matlab but will have to download an additional installation file to run the software on their local machines. The website also provides an extended tutorial.

The first step in using GelBandFitter is to load an image of a stained gel. The program will read many types of image file including BMP (bitmap), GIF (graphics interchange format), JPEG (joint photographic experts group) and Tiff (tagged image file format) formats. It will also read gray-scale, color-mapped and true-color images. Once the image has been successfully loaded, it is displayed at the bottom left of the computer screen. A screenshot of a typical GelBandFitter session is shown in Fig 1.

The user can then select a specific region of the gel for further analysis by dragging a box across the displayed image with the computer mouse. GelBandFitter calculates the mean pixel intensity for each horizontal line in the selected region and displays the resulting intensity trace in the top right of the user screen. (The averaging can significantly increase the signal to noise ratio of the intensity trace but it will also blur features if the bands are not horizontal. Images should be “pre-rotated” using conventional image-editing software if the protein bands run diagonally in the original scan.) The optical density profile is the intensity trace multiplied by -1 and offset to zero. GelBandFitter calculates this profile automatically and displays it at the bottom right of the computer screen. If so desired, the program can also subtract a trend-line from the intensity data to reduce the effects of non-uniform background staining.

Although the methods used by GelBandFitter could be extended to analyze an arbitrary number of overlapping bands, the current version of the software assumes that the user wishes to calculate the relative content of just two proteins. The software provides “Split” and “Fitting” methods for performing the calculation.

The Split method simply separates the densitometry profile into two regions with a straight line oriented at right angles to the densitometry baseline (Fig. 2C). This is equivalent to comparing the integrated pixel intensities of two adjacent regions of interest containing the upper and lower bands on the gel, respectively. (The documentation for the GelAnalyzer plugin that ships with ImageJ describes a series of user steps that complete a similar analysis.) GelBandFitter automatically determines the areas of the upper (A_{upper}) and lower (A_{lower}) sections and calculates the relative amount of protein in the lower band as $A_{\text{lower}} / (A_{\text{upper}} + A_{\text{lower}})$. The relative amount of protein in the upper band is 1.0 minus this result.

When the user first selects the Split method for a given image, GelBandFitter tries to place the demarcation line at a plausible position on the densitometry profile. The user will not always be satisfied with the automatically determined location and can over-ride it if they desire. The Split method is easy to use and easy to understand. However, simple mathematical analysis shows that it can also produce inaccurate results (Fig. 3C and D).

The Fitting method works very differently. It operates on the assumption that each protein has a densitometry profile that can be described by a mathematical function with (in the current version of the software) four parameters. GelBandFitter superposes two such functions *in silico* and then uses automatic procedures to adjust the values of the six free parameters (two are shared for both curves) until the sum of the functions is the best possible fit to the measured densitometry profile. When this iterative process finishes, GelBandFitter calculates the areas

under the two functions (A_{upper} and A_{lower}) and outputs the relative amounts of the protein in the upper and lower bands as described for the Split method in the preceding section.

The main advantage of the Fitting method is that the densitometry profile does not have to be split into separate regions of interest. This is particularly important for determining the relative content of two proteins when the densitometry profile does not drop to baseline between the bands. The main disadvantage is that non-linear multidimensional optimization (the mathematical term for the curve-fitting approach outlined above) is a difficult numerical task and there is no guarantee that the fitting procedure will converge to an appropriate profile [1].

GelBandFitter uses several approaches to minimize this problem. It (i) automatically sets appropriate starting parameter values, (ii) allows the user to choose between two different optimization strategies (a Nelder–Mead Simplex method [2] and a Broyden–Fletcher–Goldfarb–Shanno Quasi-Newton approach [3]) and (iii) allows each parameter to be individually constrained. Tests in our laboratories show that the algorithms produce appropriate fits in the vast majority of cases. Additional details on how to deal with problem situations are provided at <http://www.gelbandfitter.org>.

The user can chose to use the “Fitting” method with either Gaussian or Lorentzian curves. The Gaussian mode tries to fit the measured densitometry profile $D(x)$ with $G(x)$, the sum of two asymmetric Gaussian curves [4], where

$$G(x) = A_1 \exp\left(-\alpha(x'_1 - x_1)^2\right) + A_2 \exp\left(-\alpha(x'_2 - x_2)^2\right) \quad \text{and} \\ x'_i = x + \phi(x - x_i) \quad \text{for } x \geq x_i \quad \text{where } i=1, 2 \\ = x \quad \text{for } x < x_i \quad (1)$$

x is the distance from the top of the gel, and A_1 , A_2 , and x_1 , x_2 are parameters describing the amplitudes and positions of the two curves. α defines the curvature of the Gaussians and the skew parameter ϕ allows the functions to become asymmetric. These parameters have the same value for both curves. This corresponds to a physical situation in which the two proteins produce densitometry profiles with the same functional form. The bands differ only in their amplitude and position on the gel.

The Lorentzian fitting mode is similar but tries to fit $L(x)$ to the densitometry profile, where

$$L(x) = \frac{A_1 \gamma}{\pi(\gamma^2 + (x' - x_1)^2)} + \frac{A_2 \gamma}{\pi(\gamma^2 + (x' - x_2)^2)} \quad (2)$$

γ defines the curvature of the Lorentzians, π is the irrational constant 3.14159... and other parameters are defined as above. The asymmetric Gaussians produced the best fits in most of our tests but other researchers have suggested that Lorentzian functions are a good choice for fitting protein bands [5].

GelBandFitter was originally developed to help analyze gels showing partial separation of the α and β isoforms of mammalian cardiac myosin heavy chain (MHC). In rats, the α isoform has a predicted molecular mass of 223 379 Da (Genbank Accession No. NP_058935) while the β isoform has a predicted mass of 222 953 Da (NP_058936). These two isoforms are difficult to resolve completely by gel electrophoresis [6] and provide a practical example of the advantages of the curve-fitting technique.

Figure 2A shows ventricular samples from six different Fischer 344 rats run on a 7% acrylamide (50:1 with bis-acrylamide) resolving gel containing 35% glycerol [7]. The gel was stained with a commercial kit (Silver Stain Plus, Bio-Rad, Hercules, CA) and scanned on a conventional desktop device (V500 Photo, Epson, Long Beach, CA). Figure 2B shows an enlarged version of the pair of bands enclosed by the dashed line in Fig. 2A. Figures 2C–E show the densitometry profile of the paired bands and the relative content of the lower (β) MHC isoform calculated by the (C) Split, (D) Gaussian curve fit and (E) Lorentzian curve-fit methods, respectively. The estimates of the relative content of the β MHC isoform deduced using the Gaussian and Lorentzian curve fits were 170 and 156%, respectively of the value calculated with the Split method.

Figure 3A shows the relative contents of β -MHC calculated using the Gaussian curve-fit and Split methods for 29 ventricular samples from F344 rats of different ages. The data points would lie on the dashed line if the two analysis methods produced identical results. The solid line shows the best linear fit to the experimental data. The estimates of relative content of the β -MHC obtained using the Gaussian curve-fit method were significantly higher than the same values calculated using the Split approach ($p < 1 \times 10^{-6}$, paired t -test). The Lorentzian curve fits also produced higher estimates of the relative β -MHC content ($p < 1 \times 10^{-6}$, data not shown).

These observations suggest that the Split method systematically underestimates the amount of β -MHC protein. It is difficult to test this hypothesis experimentally (because the true amounts of α and β -MHC protein in the different samples are not known) but it is easy to investigate the issue mathematically.

Figure 3B shows a simulated optical density profile calculated from the sum of two asymmetric Gaussian curves. A horizontal line drawn at the minimum of the inter-band profile separates the trace into two distinct regions but places a disproportionate amount of the lower Gaussian in the upper area. In this case, the Split method therefore underestimates the amount of protein in the lower band.

This is not always the case. The Split method can produce reasonable answers or it can underestimate or overestimate the relative content of the bottom band. The magnitude and direction of the error depends on the separation of the bands, the relative intensity of the bands and whether the bands have asymmetric profiles.

Figures 3C and D illustrate the complexity of the situation. Simulated densitometry profiles were calculated by superposing Gaussian curves with either (C) symmetric or (D) asymmetric profiles with different inter-band separations and different relative protein contents. The Split method was then used to calculate the relative amount of protein in the lower simulated band. This method predicts the correct relative content of the lower band when the peaks are well separated (left-hand sides of each graph) but becomes inaccurate when the band profiles start to overlap. The right-hand edge of each line marks the last point at which there is an inter-band minimum in the simulated density profile.

The simulations summarized in Fig. 3D predict that when the relative content of the lower band is ~ 0.9 , the Split method will calculate the correct relative protein content irrespective of the band separation. This is in close agreement with the β -MHC content (0.853) at which the best linear fit to the experimental data in Fig. 3A intersects with the line $y = x$. This supports the hypothesis that MHC bands run (at least under the present experimental conditions) with asymmetric profiles. Curve-fitting techniques are therefore likely to provide a more accurate way of assessing relative protein content than analyses based on variants of the Split method.

In summary, estimates of relative protein content obtained by splitting the densitometry profile with straight lines can be inaccurate when the protein bands are poorly resolved. GelBandFitter

is a computer program that uses curve-fitting techniques to produce more appropriate estimates of relative protein content from images of scanned gels and immunoblots. It can be downloaded (at no cost to academic users) from <http://www.gelbandfitter.org>.

Acknowledgments

This work was supported by American Heart Association Scientist Development Grant 0630079N, NIH AG021862, NIH HL 090749 (all to K.S.C.) and by NIH HL077196 to M.L.G.

The authors M.I.M. and K.S.C. are employed at the University of Kentucky. The research grants are from NIH and the American Heart Association.

Abbreviations

A_{lower}, area of the lower section; A_{upper}, area of the upper section; >MHC, myosin heavy chain.

References

1. Press, WH.; Teukolsky, SA.; Vetterling, WT.; Flannery, BP. Numerical Recipes in C – The Art of Scientific Computing. Cambridge University Press; Cambridge: 1992. p. 394–455.
2. Lagarias JC, Reeds JA, Wright MH, Wright PE. SIAM J. Optim 1998;9:112–147.
3. Dennis, JE.; Schnabel, RB. Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Society for Industrial and Applied Mathematics; Philadelphia: 1996.
4. Vohradsky J, Panek J. Electrophoresis 1993;14:601–612. [PubMed: 8375351]
5. Shadle SE, Allen DF, Guo H, Pogozelski WK, et al. Nucleic Acids Res 1997;25:850–860. [PubMed: 9016637]
6. Warren CM, Greaser ML. Anal. Biochem 2003;320:149–151. [PubMed: 12895480]
7. Tikunov BA, Sweeney HL, Rome LC. J. Appl. Physiol 2001;90:1927–1935. [PubMed: 11299287]

Figure 1.
GelBandFitter. A screenshot showing the analysis of a typical electrophoretic doublet.

Figure 2.

Three different methods for calculating relative protein content. (A) A stained gel showing partial separation of the α and β -MHC isoforms in samples of rat myocardium. (B) Enlarged and vertically scaled view of the demarcated region in panel A. (C, D and E) Relative β -MHC content calculated using the Split method and the Gaussian and Lorentzian curve fits, respectively. A.U. stands for arbitrary units.

Figure 3.

Analysis of results calculated using the Split method. (A) Relative β -MHC content calculated by Gaussian curve fitting for 29 ventricular samples plotted against the corresponding Split method estimate. (B) Dashed lines show asymmetric Gaussian curves calculated with $\alpha = 1/2\sigma^{-2}$ and $\Phi = 1$. The solid curve is the sum of the Gaussians. (C and D) The Split method estimates of the relative content of the bottom band for simulated profiles. Different lines show calculations for different relative contents. Panel C shows calculated values for symmetric ($\Phi = 0$) Gaussians. Panel D shows calculated values for asymmetric Gaussians ($\Phi = 1$).