Molecular Weight Determination by SDS-PAGE

Introduction
Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) is a reliable method for determining the molecular weight (MW) of an unknown protein. The first step in MW determination of a protein is to separate the protein sample on the same gel with a set of MW standards. Next, a graph of log MW vs. relative migration distance (Rf) is plotted, based on the values obtained for the bands in the MW standard. The MW of the unknown protein band is then calculated by interpolation using this graph. The key to determining MW accurately is selecting separation conditions that will produce a linear relationship between log MW and migration within the likely MW range of the unknown protein.

SDS-PAGE Conditions for MW Determination
To ensure accurate MW determination, both the protein standards and the unknown protein must be electrophoresed on the same gel under identical separation conditions. It is also important to generate multiple data points (at least three gels) so that the estimated MW has statistical significance. The sample buffer used to solubilize the proteins should contain reducing agents (dithiothreitol or β-mercaptoethanol) to break disulfide bonds, which minimizes the effect of secondary structure on migration. In addition, a strong ionic detergent such as SDS is a required component of the sample buffer. SDS provides two functions: It denatures secondary, tertiary, and quaternary structures by binding to hydrophobic protein regions, and its binding confers a net negative charge on the proteins, which also results in a constant charge-to-mass ratio. The proteins are then separated through a gel in an electrical field according to their mass. However, other factors may also influence protein separation. These factors are discussed briefly under Limitations.

Analysis of Electrophoresed Proteins
An example of the approach is shown in Figure 1 using Green Fluorescent Protein (GFP) as a hypothetical example of an unknown protein. A dilution series of an E. coli lysate spiked with GFP and Precision Plus Protein™ unstained standards was electrophoresed in different lanes of a Criterion™ 4–20% SDS-PAGE gel, then stained with Bio-Safe™ Coomassie stain and destained in distilled water to visualize the protein bands. The gel was then analyzed to obtain the Rf values for each band. The Rf is defined as the migration distance of the protein through the gel divided by the migration distance of the dye front. The distance should be measured from the top of the resolving gel to the band of interest, as illustrated on the gel.

A plot of log MW versus Rf (Figure 2) was generated from the bands in the gel shown in Figure 1 to determine the MW of the unknown protein. The simplest method for this is to base the MW determination on a standard curve. If the curve is nearly linear, it can be described by the formula \( y = mx + b \), where y is the log MW, m is the slope, x is the Rf, and b is the y-intercept, as shown in Figure 2.

![Figure 1. Example showing approach for MW determination of an unknown protein. Lane 1, 10 µl of Precision Plus Protein unstained standards; lanes 2–7, a dilution series of an E. coli lysate containing a hypothetically unknown protein (GFP). Proteins were separated by SDS-PAGE in a Criterion 4–20% Tris-HCl gel and stained with Bio-Safe Coomassie stain. Gel is shown actual size. MW standards are in kD.](image)

Alternatively, Bio-Rad’s The Discovery Series™ Quantity One® 1-D analysis software, or software with similar capabilities, can be used to determine the Rf values. Quantity One is a tool for imaging and analyzing 1-D electrophoresis gels, dot blots, slot blots, and colony counts. For accurate MW determination, the unknown protein should be within the linear range of the standard curve, and the amount of the unknown protein (or its intensity after staining) should match the corresponding...
standard; in Figure 1, the unknown bands in lanes 5 and 6 are optimal. A gradient gel (for example, 4–20%) is generally used to determine the range where the unknown protein’s MW falls, because the gradient allows proteins spanning a wide MW range to be examined. Single-percentage gels can then be used to further analyze the unknown protein. For a single-percentage gel, it is important to determine the MW range in which the relationship of log MW to migration distance of the standards is linear. The accuracy of the calculated MW depends on the linearity of the relationship, represented by the \( r^2 \) value. The closer the \( r^2 \) value is to 1.0, the better the fit of the data points to a line. If a set of protein standards does not generate a linear relationship, it is acceptable to remove the data points that lie beyond the linear range, as shown in Figure 3. In Figure 3B, the values for the five largest proteins were omitted from the calculations, increasing the \( r^2 \) value from 0.913 to 0.997. Most scientific calculators and software can generate the \( r^2 \) value and the equation of the best-fit line from the data points. A comparison of results obtained using a nonlinear and a linear curve is given in the Table.

**Procedure**

Figures 1 and 2 illustrate the procedure.

1. Run the standards and unknown on an SDS-PAGE gel.
2. Process the gel with the desired stain and then destain to visualize the protein bands.
3. Determine the \( R_f \) graphically or using Quantity One software (or equivalent).
4. Use a graphing program to plot the \( R_f \) versus log MW. From the program, generate the straight line equation \( y = mx + b \), and solve for \( y \) to determine the MW of the unknown protein.

**Determining \( R_f \), Graphically**

- Use a ruler to measure the migration distance from the top of the resolving gel to each standard band and the dye front.
- Calculate the \( R_f \) value of each band using the following equation:
  \[
  R_f = \frac{\text{migration distance of the protein}}{\text{migration distance of the dye front}}
  \]

**Determining \( R_f \) With Quantity One Software**

- The \( R_f \) and \( r^2 \) values are determined automatically by the software (see Quantity One manual).

**Example of Calculation**

From Figure 1:

- Migration distance of unknown protein: 45 mm
- Migration distance of dye front: 67 mm
- So \( R_f = \frac{45 \text{ mm}}{67 \text{ mm}} = 0.67 \)

From Figure 2:

- \( y = -1.9944x + 2.7824 \)
- \( x = R_f \) of unknown protein = 0.67
- \( y = \log \text{ MW} \)
- So MW = \( 10^y = 10^{-1.9944(0.67)} + 2.7824 = 28.1 \text{ kD} \)

The actual MW of GFP is 28.3 kD, so the accuracy of the observation using these standards was 99.2%, which is below the level of error inherent in making the measurements. Note that calculation based on the nonlinear curve (see Table) would not have been as accurate; if values had been calculated based on the entire set of standards shown in Figure 3, the predicted value for GFP would have been 37.0 kD.

**Limitations**

MW determination by SDS-PAGE is a dependable method. However, an unknown protein’s MW should always be obtained by mass spectrometry if a more precise MW determination is needed. Mass spectrometry has a higher degree of accuracy because each amino acid of a protein is analyzed. Protein-to-protein variation can be minimized by denaturing samples, reducing proteins, normalizing the charge-to-mass ratio, and electrophoresing under set conditions. However, factors such as protein structure, posttranslational modifications, and amino acid composition are variables that are difficult or impossible to minimize and can affect the electrophoretic migration.

A few examples demonstrate the effect of these variables. Glycoproteins migrate unpredictably in SDS-PAGE (Hames 1998). The hydrophilic glycan moieties can obstruct the binding of SDS, and the decreased hydrophobic interaction between the protein and SDS result in an inconsistent charge-to-mass ratio. However, some evidence suggests that glycoproteins exhibit more normal protein migration in gradient gels. Acidic proteins (such as tropomyosin) also migrate abnormally on SDS-PAGE gels. The acidic residues may be repelled by the negatively charged SDS, leading to an unusual mass-to-charge ratio and migration. Highly basic proteins...
(for example, lysozyme, histones, and troponin I), which contain an abundance of positively charged amino acids, migrate more slowly in SDS-PAGE due to a reduced charge-to-mass ratio, resulting in a higher apparent MW. Proteins with high proline content or with other unusual amino acid sequences (for example, ventricular myosin light chain) show a decreased electrophoretic mobility as a result of kinks and structural rigidity caused by the primary sequence. These differences can contribute to an error of ±10% when using SDS-PAGE to determine the MW of a protein. Following SDS-PAGE analysis with mass spectrometry will produce a more accurate determination. Despite these limitations, SDS-PAGE is still a commonly used method for MW determination of a protein.

Reference

Table. Comparison of results using nonlinear and linear ranges of standards to determine MW of an unknown protein. The protein was run on three Criterion 15% SDS-PAGE gels with Precision Plus Protein unstained standards to determine its MW. The accuracy of the MW determination is dependent on the linearity of the curve; curves used were those in Figure 3. The hypothetically unknown protein was GFP.

<table>
<thead>
<tr>
<th></th>
<th>Nonlinear Curve</th>
<th>Linear Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>MW determination*</td>
<td>37.0 kD</td>
<td>29.3 kD</td>
</tr>
<tr>
<td>Accuracy of MW estimate</td>
<td>69.3%</td>
<td>96.5%</td>
</tr>
</tbody>
</table>

* Actual MW is 28.3 kD based on amino acid sequence.

Fig. 3. Comparison of nonlinear and linear curves from a 15% Criterion SDS-PAGE gel. A, a plot of the log MW versus Rf generated using the entire protein standard range demonstrates a nonlinear curve, resulting in inaccurate MW determination. B, a curve generated from the proteins in the linear range (10–37 kD). The extreme points were removed, producing a linear curve and thus an accurate MW determination.
Ordering Information

Description
Criterion Tris-HCl Gels

170-4068 Criterion Blotter With Plate Electrodes, includes cell assembled with plate electrodes, lid with cables, 3 sample loading guides (12+2 well, 18-well, 26-well), instructions

170-4070 Criterion Blotter With Wire Electrodes, includes cell assembled with wire electrodes

170-6461 Immun-Blot® Goat Anti-Rabbit IgG (H + L)-AP Kit

170-6464 Immun-Blot® Goat Anti-Mouse IgG (H + L)-HRP Kit

170-6431 HRP Conjugate Substrate Kit

162-0217 Sequence-Blot PVDF/Filter Paper Sandwiches, 7 x 8.5 cm, 20 pack

162-0212 Nitrocellulose/Filter Paper Sandwiches, 0.2 µm, 7 x 8.5 cm, 20 pack

162-0214 Nitrocellulose/Filter Paper Sandwiches, 0.45 µm, 7 x 8.5 cm, 20 pack

162-0215 Nitrocellulose/Filter Paper Sandwiches, 0.45 µm, 7 x 8.5 cm, 50 pack

Blotting Membrane/Filter Paper Sandwiches (7 x 8.5 cm)**

162-0233 Nitrocellulose/Filter Paper Sandwiches, 0.2 µm, 8.5 x 13.5 cm, 50 pack

162-0232 Nitrocellulose/Filter Paper Sandwiches, 0.2 µm, 8.5 x 13.5 cm, 20 pack

162-0234 Nitrocellulose/Filter Paper Sandwiches, 0.45 µm, 8.5 x 13.5 cm, 20 pack

162-0235 Nitrocellulose/Filter Paper Sandwiches, 0.45 µm, 8.5 x 13.5 cm, 50 pack

Blotting Membrane/Filter Paper Sandwiches (8.5 x 13.5 cm)**

162-0236 Sequence-Blot PVDF/Filter Paper Sandwiches, 8.5 x 13.5 cm, 20 pack

162-0237 Sequence-Blot PVDF/Filter Paper Sandwiches, 8.5 x 13.5 cm, 50 pack

Equipment

165-6001 Criterion Cell, includes electrophoresis buffer tank, lid with cables, 3 sample loading guides (12+2 well, 18-well, 26-well), instructions

For more ordering information, visit us on the Web at discover.bio-rad.com

Coomassie is a trademark of BASF Aktiengesellschaft. Streptactin is a trademark of Institut für Bioanalytik GmbH. O/C is licensed by German patent application P 19641876.3. Bio-Rad Laboratories, Inc. is licensed by Institut für Bioanalytik GmbH to sell this product for research use only. Tween is a trademark of ICI Americas Inc.

Information in this tech note was current as of the date of writing (2004) and not necessarily the date this version (Rev A, 2004) was published.