White Paper

Linear Range Determination in Empiria Studio® Software

LI-COR®

Published September 2018. Revised December 2019.
The most recent version of this document is posted at licor.com/bio/support.

Table of Contents

I. Introduction

Validation is a critical aspect of quantitative Western blot (QWB) analysis. In addition to antibody validation, it is important to determine the relationship between sample loading and signal intensity to ensure that signals are detected within the linear range of the assay. In 2015, editors of the Journal of Biological Chemistry expressed concern about the linear range of detection for QWB experiments, indicating that this relationship should be confirmed for every antigen (Fosang, AJ and Colbran RJ. *Transparency is the key to quality.* J Biol Chem. 290:29692-694, 2015). The editors also cautioned researchers about the limited linear range of certain detection methods.

To address these issues, Empiria Studio software provides dedicated workflows to guide researchers through validation of the linear range and documentation of the results. This guidance helps the user identify the range of sample loading that produces a linear and proportional response in signal intensity for a target protein and internal loading control. To understand why the linear range of detection is so critical for QWB analysis, it is helpful to define a few important terms and describe their significance in this context.

The dynamic range of a detector or sensor is the ratio of the largest and smallest possible signal values it can measure. A very large dynamic range is needed to avoid saturation of strong signals and provide the wide linear range required for QWB analysis. The term “linearity” is often used to describe the portion of the dynamic range where the output measured by the detector is a linear function of the signal input. This relationship between input and output can be represented by a straight line. However, because this line may or

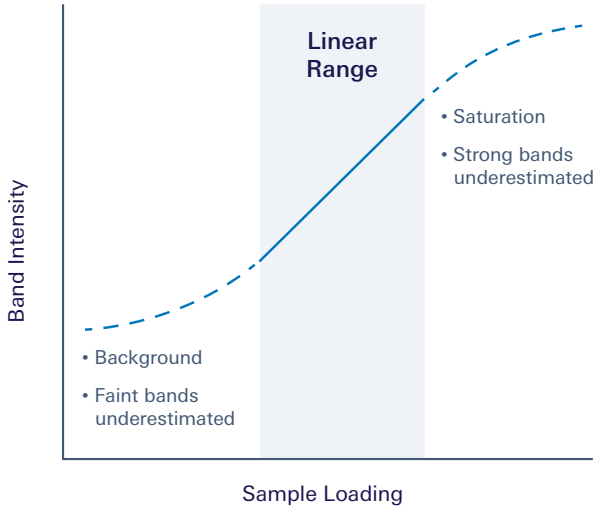
may not pass through the origin, the recorded output is not necessarily proportional to signal input.

II. Linear Range of Detection

In QWB analysis, both linearity and proportionality are important. The linear range is that portion of the dynamic range where a change in signal input produces a proportional change in the output recorded by the detector. In this range, signal input and output are related by a constant. This linear and proportional relationship is typically represented by a straight line that passes through the origin.

Linearity and proportionality determine the linear range of detection for QWB analysis. The linear range of detection is the range of sample loading where signal intensity increases in proportion to protein abundance. This range of sample loading amounts produces a linear and proportional relationship between the amount of target protein on the membrane and the band intensity recorded by the detector. Quantitative analysis must be performed within this range (Fig. 1). Above the linear range, strong signals exceed or "saturate" the capacity of the detector and produce less than the expected increase in signal intensity; below the range, faint bands are difficult to discriminate from image background.

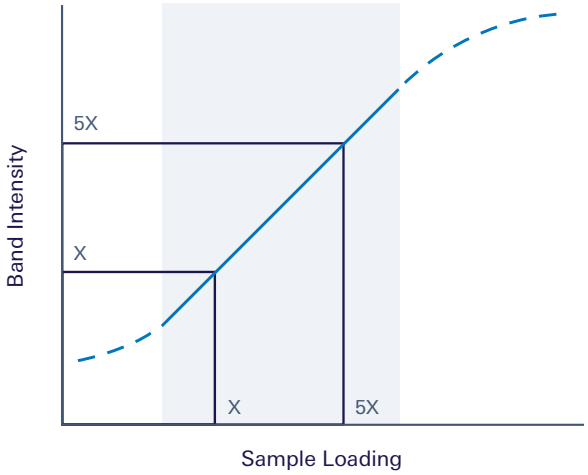
Figure 1. Quantification must be performed within the linear range of detection. Outside this range, band intensities are inaccurate. Above this range, strong bands exceed or "saturate" the capacity of your assay and do not produce the expected increase in signal intensity. The intensity of strong bands is underestimated. Below this range, faint bands are similar in intensity to the background noise and band intensity may be underestimated.



In the linear range of detection, an approximately constant ratio is observed between sample abundance and signal intensity. A change in sample loading or target protein abundance on your blot should produce a roughly equivalent signal response on your image. For example, a five-fold increase in sample loading or target protein abundance should generate an approximately equivalent increase of about five-fold in band intensity (Fig. 2). This relationship and constant ratio are the foundation of QWB analysis, and all quantification must occur within this range. If too much sample is loaded, saturation occurs and the proportional relationship between target abundance and band intensity is lost; quantification

is no longer valid. Saturated bands do not accurately represent sample abundance and cannot be used for QWB analysis.

Figure 2. In the linear range of detection, signal intensity is proportional to sample abundance. A five-fold increase in sample loading or target protein abundance should generate an approximately equivalent increase in band intensity.



III. Combined Linear Range

To analyze a target protein on a blot, the linear range of detection for that target must be determined. Typically, serial dilutions of biological sample are examined by Western blotting to empirically identify the range of sample loading that produces a linear and proportional signal response without saturation.

For QWB analysis, an internal loading control (ILC) should be included on each blot to enable normalization of the results. Normalization mathematically corrects for unavoidable sample-to-sample and lane-to-lane variation by comparing the target protein to the ILC. An ILC is an endogenous protein that is present in all samples at a stable level and can be used as an indicator of sample protein loading. Commonly used ILCs include total protein staining of membranes, “housekeeping proteins” (such as actin, tubulin, or GAPDH), and analysis of protein modifications with pan- and modification-specific primary antibodies. To provide an accurate readout of sample loading, band intensity must be proportional to the abundance of the ILC. Like the target protein, analysis of the ILC must be also performed within its linear range of detection.

QWB analysis is accurate only if the target protein and ILC signals can both be imaged within the same linear range of detection. This combined linear range of detection is the range of sample loading that produces a linear and proportional response in signal intensity for both the target and internal loading control (see Fig. 3). That range must be identified experimentally for each target and ILC. Typically, serial dilutions of protein sample are analyzed to separately determine the linear range of detection for the target and ILC. These ranges are then compared to identify the available range of sample loading where both target and ILC demonstrate a linear and proportional signal response. Whenever an experimental

parameter is changed (such as sample type, treatment conditions, transfer method, membrane, antibody, or detection reagents), the combined linear range must be re-validated.

IV. Finding the Combined Linear Range with Empiria Studio

Currently, identification of the combined linear range of detection is a manual and inherently subjective process. Empiria Studio takes an entirely new approach to this process. When Empiria Studio software is used to analyze a blot with serial dilutions of sample, it automatically plots the data and provides a combined graph that displays the target and ILC results together. A color bar above the graph indicates the regions of linearity and proportionality identified by the algorithm.

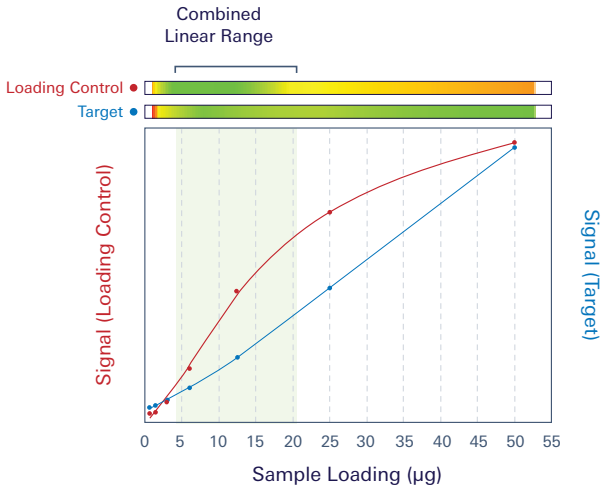
Analysis of linearity typically uses a regression to fit the entire data set to the equation of a straight line, $y = mx + b$, where b indicates the y-intercept on the graph. Linearity occurs in the range where the relationship between signal input and output can be represented by a straight line – but it does not imply or guarantee proportionality. For this reason, the R^2 value may provide little useful information for QWB analysis. R^2, the coefficient of determination, describes the variation and indicates how well that regression line represents the actual data. If the data fit the regression line perfectly, $R^2 = 1$. However, if the line does not pass through the origin, R^2 may be equal to 1 for a data set that does not display a proportional signal response.

Accurate QWB analysis requires both linearity and proportionality. A proportional relationship is defined by the equation $y = mx$, representing a straight line that passes through the origin with b (the y-intercept) equal to zero. When you fit an entire data set to a straight line, the fitted line may or may not indicate a linear and proportional relationship. The line may have a non-zero intercept ($b \neq 0$). However, even if the full data set cannot be fit to a line passing through the origin with $b = 0$, some regions of the data may be proportional.

For these reasons, Empiria Studio does not use a linear regression or R^2 value to determine the linear range. Rather than fitting a single function to define the line, multiple functions are derived in a piecewise manner to examine segments of the data set and identify regions of linearity and proportionality. In these regions, the y-intercept approaches zero. Signal intensity is proportional to the abundance of target protein or ILC and quantitative analysis can be performed.

On the View Linear Range and Combined Linear Range graphs in Empiria Studio, these regions of proportionality are indicated by green areas (excellent linearity) and yellow areas (good linearity) of the color bar. This guidance indicates appropriate ranges of sample loading in the linear range of detection. The upper color bar shows regions of proportionality for the ILC and the lower bar displays information for the target protein. Overlapping areas of green or yellow on the color bars indicate the available range of sample loading for accurate detection of the target and ILC in the combined linear range. The middle of this range is suggested as a starting point for sample loading.

Figure 3. Viewing the combined linear range graph in Empiria Studio. Overlapping areas of green and yellow in the color bars indicate the available range of sample loading for accurate detection of the target and ILC. The middle of this range should be an appropriate starting point for sample loading. In this example, that midpoint is approximately 10 µg of sample protein.



V. Using the Combined Linear Range in QWB Analysis

With the Linear Range Validation workflow in Empiria Studio, a Combined Linear Range graph is automatically generated. The slider tool above the color bars can be moved to designate the approximate boundaries of the linear range and identify an appropriate range of sample loading for accurate detection. The middle of this range is a good starting point for sample loading in a QWB experiment. If the experimental treatment is known to cause a substantial increase or decrease in target abundance, sample loading may need to be adjusted accordingly.

The extent of overlap and range of sample loading in the combined linear range will be determined by the relative abundance of the target and ILC. The combined linear range may be very narrow, or even non-existent, for detection of two proteins of very different abundance. This is important information, given that many labs routinely load a set amount of sample protein per lane, generally 10 - 50 µg. In many cases, sample loading in excess of 5 - 10 µg of total protein per lane will lead to saturation of strong signals.

Housekeeping Protein

The impact of saturation on combined linear range is often observed when an HKP is used as the ILC for normalization. Because the HKP is frequently expressed at a much higher level than the target protein, the combined linear range may be extremely limited. However, this situation makes it even more important to ensure that both proteins are detected within the combined linear range. If too much sample is loaded, HKP signals will become saturated. Band intensity will not be proportional to sample abundance, and the intensity of these strong bands will be underestimated. Saturation causes strong bands to appear similarly dark, concealing actual differences between samples. If differences in sample loading cannot be detected, accurate normalization and quantitative analysis are not possible.

Figure 4 shows a typical combined linear range result for a target protein and highly expressed HKP. Because the HKP signals quickly become saturated, the appropriate range of sample loading is limited (about 2 - 7 µg of sample protein for excellent linearity, with a midpoint of 4.5 µg). If 15 µg of sample protein were loaded for a QWB experiment (Fig. 4), the HKP would not be detected within the linear range and an increase in sample abundance would not produce the expected increase in band intensity. If a combined linear range cannot be identified, it may be necessary to choose a different ILC for the experiment.

Validation is critical when an HKP is used as the loading control for QWB normalization. Because the combined linear range is dependent on the relative abundance of each HKP and target in the sample, it must be determined for each new combination of HKP and target. It is also important to demonstrate that the HKP is stably expressed in all experimental samples and is not affected by the experimental conditions and treatments applied. If the HKP is not expressed at a stable level, it should not be used for normalization.

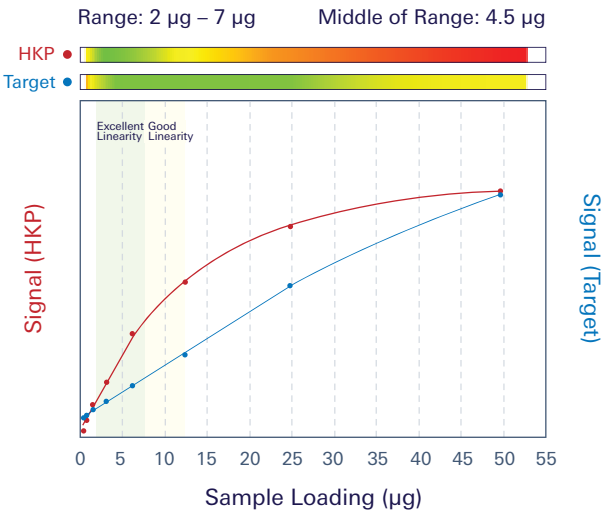


Figure 4. The combined linear range may be very narrow if a highly abundant HKP is used as the internal loading control. In this example, HKP signals quickly become saturated and are not proportional to sample abundance. Only 15 µg sample protein per lane (arrowhead) would represent significant overloading, and the resulting HKP signals would fall above the linear range. For this combination of HKP and target, the combined linear range is about 2 - 7 µg of total sample protein per lane (with excellent linearity). The midpoint of this range, and suggested starting point for sample loading, is about 4.5 µg per lane.

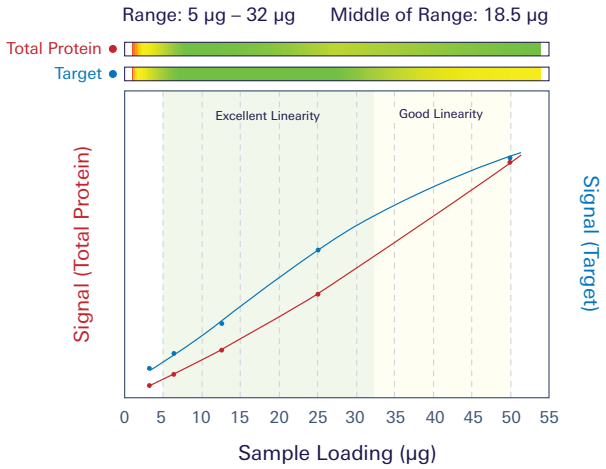
Total Protein Staining

The limitations of HKPs as loading controls have prompted interest in total protein staining as an ILC for normalization. This method is emerging as the preferred approach for normalization of sample loading in QWB analysis. After transfer from gel to membrane, a fluorescent total protein stain is used to detect the total amount of sample protein in each lane. This approach provides an aggregate measurement of many different sample proteins in each lane, to enable accurate assessment of sample loading.

Total protein staining can provide a proportional response across a much broader range of sample loading than many commonly used HKPs. This typically provides more options for sample loading in the combined linear range, making it easier to accurately detect the target and ILC on the same blot.

This effect is illustrated in Figure 5, which shows a representative combined linear range result for a target protein and total protein staining. The total protein stain demonstrates a wide linear range, without signal saturation (about 5 - 50 µg of sample protein). When overlaid with the target protein results, the combined linear range encompasses a broader range of sample loading (about 5 - 32 µg sample protein in this example with excellent linearity, with a midpoint of 18.5 µg). A wider range of appropriate sample loading provides more flexibility, for reliable analysis without saturated signals that could undermine the accuracy of the results. Unlike an HKP, total protein staining does not require validation of stable expression for each new set of experimental conditions and treatments.

Figure 5. Total protein staining may enable a wider combined linear range for detection and normalization. On this graph, the total protein stain provides a proportional response across a range of 5 - 50 µg sample protein per lane. The combined linear range for this combination of ILC and target is about 5 - 32 µg total sample protein per lane for excellent linearity, with a midpoint of 18.5 µg sample protein.



VI. Discussion

With its dedicated workflows, Empiria Studio guides the user through the linear range validation process and documents the results for future use. Validation of the relationship between sample loading and signal intensity is required to verify that both the target protein and ILC are detected in the combined linear range. All QWB analysis must be performed within this range, where signal intensity is proportional to sample abundance. The tools and guidance provided by Empiria Studio help researchers identify and document this range, for more accurate normalization and more reliable QWB results.

LI-COR Biosciences

4647 Superior Street
Lincoln, NE 68504
Phone: +1-402-467-0700
Toll free: 800-645-4267
biosales@licor.com
licor.com/bio

Regional Offices

LI-COR Biosciences GmbH

Siemensstraße 25A
61352 Bad Homburg
Germany
Phone: +49 (0) 6172 17 17 771
bio-eu@licor.com

LI-COR Biosciences UK Ltd.

St. John's Innovation Centre
Cowley Road • Cambridge
CB4 0WS • United Kingdom
Phone: +44 (0) 1223 422104
bio-eu@licor.com

Doc # 979-18531
12/19