Western Blot Normalization:
Challenges and Considerations for Quantitative Analysis

LI-COR Biosciences • 4647 Superior Street • Lincoln, NE 68504

Table of Contents

LI-COR

1. Introduction

Researchers rely on Western blotting (also called immunoblotting) to detect and compare target proteins in complex samples. This trusted technique has become a widely-used tool for quantitative analysis and comparison of relative protein levels.

Quantitative immunoblotting relies on internal loading controls to confirm that changes observed in target protein abundance represent actual differences between the protein samples. Internal loading controls verify that samples are uniformly loaded across the gel, confirm consistent transfer from gel to membrane, and enable comparison of relative protein levels among samples.

Sample-to-sample variability is inevitable, even when sample concentrations are adjusted for uniform loading. Internal loading controls can mathematically compensate and correct for this variation. This process is called normalization, and is used to improve the accuracy and reproducibility of quantitative Western blot analysis. Careful, appropriate normalization is critical for meaningful comparison of relative protein levels and responses, and is particularly crucial when small or subtle responses are reported.[1-3] Accurate normalization begins with careful experimental design, to reduce the amount of sample-to-sample correction required. Reducing or eliminating sources of error will improve sample reproducibility and minimize the contribution of normalization to data analysis.

Internal loading controls are an important part of immunoblotting studies, and are often required for publication. But what's the best way to use internal controls for normalization and quantitative analysis? Many researchers who use a "housekeeping protein" (such as actin or tubulin) as an internal loading control assume that this method is reliable, because it's widely published. But recent studies indicate that housekeeping proteins are sometimes less reliable than expected as loading controls.[4-9]

This raises a number of questions. What's the best normalization strategy? How do loading controls affect the statistical properties of the data?[10] Is it better to stick with what we know, or explore different ways of thinking? As journals change their reporting requirements, what will editors and reviewers expect?

These questions circle back to the role of internal loading controls: confirming that the changes you see on the blot reflect actual change in the composition and/or biology of your samples. To demonstrate statistically significant changes in the abundance of target protein, you need a reliable normalization strategy that fits the context and biology of your experiment. But normalization can be more complex than it appears. It's important to understand the limits of each normalization approach – because your current method may not be the best fit for your experimental conditions.

This paper describes important considerations, strengths, and limitations of commonly used loading controls and normalization methods, to assist you in choosing a reliable method for your experimental context. Alternative approaches to Western blot normalization are examined, and the effects of normalization strategy on resulting data are explored. Important considerations for using a single internal loading control protein are discussed, as well as emerging alternatives like total protein staining that use aggregate measurements for normalization of sample loading.

2. Understanding Western blot normalization

2.1 How does normalization work?

Quantitative immunoblotting is often used for ratiometric analysis, to compare relative protein levels across a group of samples. Relative comparison requires that ALL samples be uniformly loaded with equal amounts of sample protein, without variation; ratiometric analysis is based on this fundamental assumption. But complete uniformity is a flawed assumption. A variety of factors can introduce unavoidable sample variability and inconsistency across each blot.[11]

In quantitative immunoblot analysis, a target protein is measured in samples from various experimental conditions. The intensity of each target protein band is then divided by the intensity of the internal loading control for that sample. This adjusts target protein signals with respect to small, unavoidable variations in cell number and sample loading.[10] The ratio of target protein to loading control is then used to compare target protein abundance in different samples (Fig. 1).

The concept of normalization assumes that both measures (target protein and loading control) are dependent on sample concentration, and they will vary together to the same degree.[11] If this is true, calculating the ratio of target protein to internal control (i.e., dividing or "normalizing" by the loading control) will correct for variability. Accuracy of the normalization process is critical for quantitative analysis and meaningful comparison of protein samples.

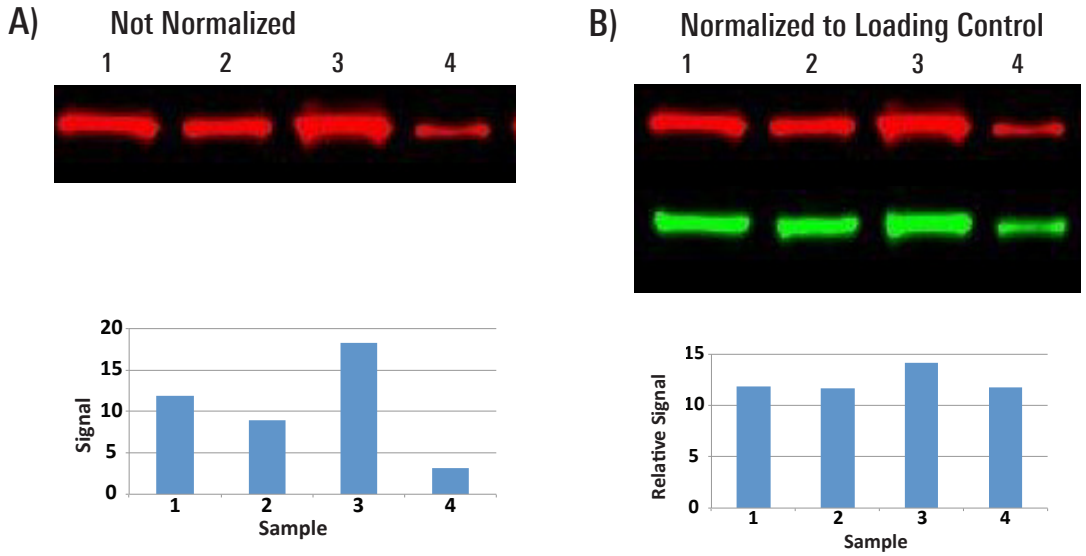

Figure 1. Normalization makes quantitative analysis more accurate. Internal loading controls and normalization ensure that observed changes in protein levels represent actual change in protein samples, not experimental artifacts. A) Raw data (not normalized). Target protein levels (red) are variable, with sample 4 showing the lowest protein level. B) Normalized data. The internal control protein (green) reveals that protein loading was inconsistent. The normalized data demonstrate that target protein levels are very similar in all samples.

Reducing or eliminating sources of variability throughout the experiment will improve sample reproducibility. This minimizes the amount of correction applied by normalization, enhancing the precision and reproducibility of sample-to-sample comparisons. Normalization is intended to compensate for small and unavoidable variation between samples, and cannot completely remove variability. If samples are inconsistent or large data corrections are applied, accuracy may be affected.

Precision and accuracy are even more crucial for interpretation of small changes in protein expression.[3] Reproducible measurement of a small response demands a much higher level of precision than a strong response, because the magnitude of the response must exceed the variability of the assay. If an immunoblot assay displays 20 - 30% variability (often cited as a typical coefficient of variation for Western blotting[1,11-12]), the measured response must exceed that variation by a statistically significant margin. If assay variability is low, it may be possible to reproducibly document a subtle response. Effective, carefully-planned normalization will more accurately reflect the amount of protein in each lane, and the actual composition of the samples.

2.2 Reporting normalization in published studies

Normalization is very relevant to the ongoing conversation about reproducibility in science. Although normalization is critical for quantitative immunoblotting, published studies frequently omit the experimental details needed for other researchers to evaluate the results.[13-18] Fortunately, scientific publishers are now increasing transparency – asking authors to report the "routine" details of their experimental methods, including data analysis methods such as immunoblot normalization.

In response to concerns about reproducibility, *Nature* and the *Nature* research journals introduced new editorial measures.[18] The updated guidelines ask authors to "describe methodological parameters that can introduce bias or influence robustness, and provide precise characterization of key reagents that may be subject to biological variability…[as well as] more precise descriptions of statistics". Space restrictions in the methods section were abolished to encourage detailed descriptions of experimental design and methods.[18] As more journals move in this direction, other publishers, and even funding agencies, may follow suit.

3. Common loading controls and normalization methods

A variety of biological factors can introduce error in Western blot analysis. The cell number may change in response to cell death or stimuli. Confluency of cultured cells drives considerable variability in protein expression.[19] Structural and metabolic proteins thought to have stable expression may be affected by growth conditions, stimuli, cell type, developmental stage, and other factors.[4-9] Biological context can and should influence the selection of appropriate internal loading controls and normalization strategies. It's not only important to have an internal control, but also to understand why that control is appropriate for the context and biology of the experiment.

An effective loading control meets two requirements: its expression is relatively constant across conditions and samples relevant to the experiment, and the resulting signal intensity is a linear representation of loading control abundance. If a single internal loading control cannot meet both requirements, more than one loading control should be used in that experiment.[1-3,11] As evidence emerges that single loading controls may not consistently meet both criteria, aggregate methods of normalization for quantitative immunoblot analysis are becoming popular alternatives.

A recent study reframed commonly-used normalization strategies and evaluated their statistical impact on data interpretation. Although this topic has been examined in depth for microarrays, the effects of normalization on the coefficient of variation (CV) and interpretation of Western blot data aren't well understood. For each method, the authors explored the effects of loading control variability on the normalized data.

Common immunoblot normalization strategies can be divided into two categories: normalization by a fixed point (using a single internal control protein) and normalization by sum (using total protein staining or multiple internal controls). The merits, liabilities, and statistical impact of each method are discussed below.

3.1 Normalization by a fixed point: Using an internal loading control protein

Normalization by a fixed point (using a single internal control protein) is the most commonly published method.[10] A second, unrelated protein that is common to all samples (often a "housekeeping protein" such as actin, tubulin, or GAPDH) is typically detected on the same blot as the target protein. Expression of the internal reference protein is assumed to be stable in all samples and unaffected by experimental conditions, such that reference protein abundance is entirely dependent on sample concentration.

This approach divides the data for each sample by measurement of a single reference point in that sample, the internal control protein. Biological variability of that fixed point can introduce uncertainty and increase the mean CV of the normalized data. The extent of influence depends on the choice of control for the experiment; internal controls with higher variability will more strongly affect data analysis. Normalization by a fixed point reduces false positives in the normalized data (samples with no statistically-significant difference in protein level that are mistakenly identified as different in the data analysis). But fixed-point analysis can also greatly increase the percentage of false negatives (samples that have small but statistically-significant changes in protein level that are not identified by data analysis). As a result, this method may sometimes fail to identify actual differences between samples.[10]

Using a single internal control protein alters the Western blot experiment. Rather than measuring the level of target protein relative to a sample characteristic such as total protein or cell number, you are measuring target protein levels relative to the protein selected as the internal control.[11] That protein may be very appropriate, if it is stably expressed and its abundance in all samples is dependent only on sample concentration.

Recent studies, however, indicate that stable expression of housekeeping proteins shouldn't always be taken for granted.[4-9] Although their expression is often thought to be relatively stable, steady-state levels of these proteins are known to vary in different tissues, or to be altered by some experimental conditions (discussed in section 7.1).[4-9] If a single internal reference protein will be used for normalization, stable expression should be validated with the appropriate sample types, experimental conditions, and antibodies.[1-9] This issue is explored in section 6, and several published examples are presented.

Band intensity and signal saturation also affect the accuracy of a single internal loading control. The intensity of internal control bands should fall in the middle of the linear range of detection whenever possible.[10] But many proteins used as internal controls, such as actin and tubulin, are highly expressed and generate strong bands – often falling outside the linear range of detection and becoming saturated.[1-3,11,13] As a result, signal intensity cannot increase in proportion to protein abundance. High-intensity normalization points increase the mean CV of the normalized data and should be avoided.[10]

Saturation artifacts are very deceptive – they can mask actual variation in protein levels by underestimating the amount of protein in the saturated bands. The similar intensities of saturated bands create the false impression that protein levels are nearly equivalent (Fig. 2).

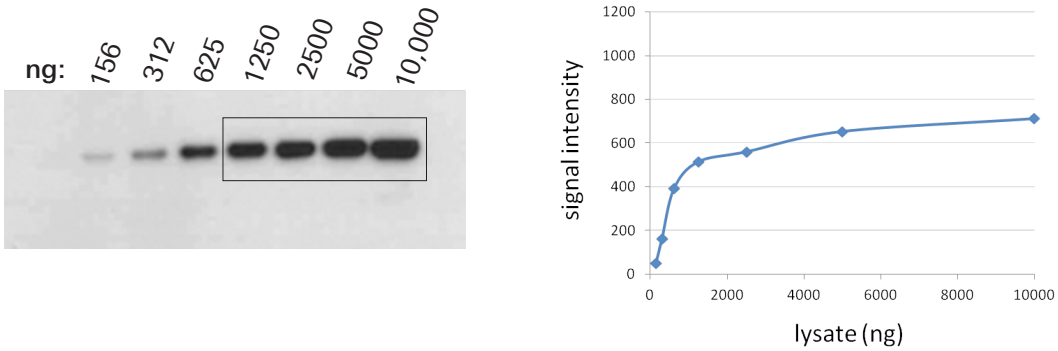

Figure 2. **Strong bands become saturated and underestimate protein abundance.** Target protein was detected in serial dilutions of NIH-3T3 cell lysate, using chemiluminescent substrate and 5-min film exposure. Strong signals (box) display saturation because they fall outside the linear range of detection. Band intensity can no longer increase proportionately to indicate protein abundance. As a result, the signal intensity of the saturated bands appears similar. High-intensity data points should not be used as controls for normalization.

3.2 Normalization by sum: Using total protein staining or multiple internal loading controls

Some researchers recommend the use of multiple internal control proteins on each blot.[1,11] This is a form of normalization by sum, an aggregate method for analysis of protein levels. Normalization by sum, using a multi-protein loading control strategy and aggregate measurements, is an emerging alternative to the fixed-point normalization strategy. In this approach, the target protein signal for each sample is divided by the sum of data obtained for that sample using total protein staining or multiple internal control proteins.

Unlike the fixed-point strategy, normalization by sum generally does not introduce uncertainty. It redistributes the uncertainty of the raw data in a mean-dependent manner, reducing the variability of high-intensity measurements and increasing the variability of low-intensity measurements.[10] This normalization strategy may increase false positives for high-intensity data points.

3.2.1 Total protein staining

Total protein staining is one type of normalization by sum. This approach uses a stain or chemical label to detect the total amount of sample protein and correct for variation. Total protein staining measures the aggregate protein signal (sum) in each lane, and eliminates the error that can be introduced by a single internal control protein.[3,5,9-11]

Total protein staining is emerging as a reliable and widely applicable strategy for quantitative immunoblotting.[2,11] It directly monitors and compares the aggregate amount of sample protein in each lane, rather than using an internal reference protein as a surrogate marker of sample concentration. This direct, straightforward approach to protein quantification may increase the accuracy of normalization. And it may also simplify your normalization workflow – because, unlike an internal control protein, total protein staining does not require validation for each experimental context and biological system.

An ideal total protein stain should meet several key requirements. It should produce linear signal output in response to sample concentration, across a wide range; correct for variation at all points in the Western blot process, including gel loading and transfer to membrane; and be compatible

with downstream immunodetection of target proteins.[1,11,20-22] The linear range of detection should be determined empirically to ensure that signal intensity is moderate, without saturation or low signal-to-noise ratios.[10]

Total protein staining is performed by gel staining or by staining of the membrane after transfer. Alternative methods for total protein detection, such as the Bio-Rad Stain-Free™ technology, replace conventional staining methods with chemical labeling of the sample proteins before or during electrophoresis. These methods are discussed in section 7.

3.2.2 Using multiple internal control proteins on a single blot

Multiple internal control proteins represent another approach to normalization in sum. In this strategy, several internal control proteins are detected on the same blot, to provide a more thorough assessment of sample loading. The band intensities of multiple loading controls are aggregated and used to calculate a mean estimate of sample concentration. This aggregate strategy is less sensitive to biological variability than a single loading control, but may require optimization.

A detailed 2015 study tested this approach, analyzing a single target protein in multiple cell lysate samples (biological replicates) using a fluorescent immunoblot method.[1] An antibody cocktail was also used to simultaneously detect five internal control proteins in each lane of the blot. After imaging, data analysis examined the effect of each loading control, as well as various combinations of loading controls, on reproducibility in these biological replicates. Without normalization, the coefficient of variation (CV) was 21%. Normalization with a single loading control produced variable CV results, depending on the protein used. GAPDH normalization decreased the CV to 9%, whereas normalization with tubulin did not improve the CV. The poor performance of tubulin in this experiment is likely due to saturation effects.

When additional loading controls were included in the analysis, the CV steadily improved to 7-8%. With more loading controls, reproducibility became less dependent on the specific combination of controls used for analysis. Because multiple controls are averaged, saturation or variability of one internal control has a negligible effect on the overall accuracy of normalization. This study concludes that using a single internal control protein for normalization can be risky, and recommends a multi-protein normalization strategy for quantitative immunoblot analysis. Three or more internal loading controls are suggested, if the sample concentration is unknown or cannot be directly measured.[1]

4. Important factors for effective normalization

4.1 Experimental design

Accurate normalization begins long before the gel is loaded, by carefully planning the experiment to make normalization as unnecessary as possible.[3,11] Although normalization is a powerful method, it cannot completely remove variation and should be used judiciously.[11]

Throughout the experiment, methodological choices and seemingly minor variations will affect the reproducibility and linearity of your immunoblot results. Without careful attention, immunoblotting may produce pseudo-quantitative results that are not proportional to the input.[1-3,20] A thoughtfully planned experimental strategy can control and reduce error, maximizing the consistency and reproducibility of the resulting samples. This, in turn, minimizes both the amount of correction required and the contribution of normalization to your data analysis.

A recent "Research Resource" article in *Science Signaling* proposed simple Western blot diagnostics to identify and optimize a variety of factors that can affect your results.[1,16] Several of these factors are briefly described here.

4.2 Sample preparation

Sample preparation can profoundly affect experimental outcome.[1-3] Even minor inconsistencies in plating, cell lysis, reagent volume, and other technical details can have a surprising impact.

Cell lysis conditions can affect sample composition, because they strongly influence protein extraction, solubilization, and modification status. Relevant proteins may be lost in the insoluble fraction; experimental treatments or stimuli may cause a shift between soluble and insoluble fractions.[1]

After sample preparation, a protein assay should be used to estimate the total protein concentration of each sample. Bradford, BCA, and Lowry assays are widely used; Nanodrop assays (A_{280}) are an option for cytosolic samples with no nucleic acid contamination.[23] The choice of protein assay should be influenced by the presence of detergents, buffers, or other components of your samples. Gel loading should be adjusted according to the estimated protein concentration, and samples loaded as accurately as possible to reduce the need for normalization.[1,10] Some researchers run a "pilot gel" and use Coomassie staining to confirm the results of the protein assay and fine-tune sample loading as needed.

4.3 Sample loading

Sample loading is a critical but underappreciated parameter. Many labs routinely load a specific, set amount of total protein per lane, generally 10-50 µg.[1-3,11] Sample loading is rarely optimized, yet overloaded gels are a significant problem.[24] If samples are overloaded, you may see very consistent band densities for an abundant internal loading control protein. However, this deceptive similarity may occur because those strong bands have exceeded the local capacity of the transfer membrane and/or exceeded the linear dynamic range of detection (signal saturation is discussed in detail in section 5). Running a standard curve with two-fold serial dilutions of cell lysate will identify the linear range and approximate point of saturation, which may be different for each target protein (Fig. 2).[3]

Janes examined this parameter using two-fold serial dilutions, from massive overloading (200 µg cell extract) to below the limit of detection (100 ng extract).[1] The linear range of sample loading was then determined for several target proteins. For p38 and actin, linear detection was demonstrated up to 50 µg of total protein. But for Hsp90 and tubulin, saturation was observed with less than 25 µg of total protein. Gomes et al[2] indicate that 20 µg/lane is excessive; Taylor and Posch report that sample loading in excess of 5-10 µg/lane are frequently overloaded and saturated.[3]

4.4 Detecting multiple proteins on the same blot

Quantitative immunoblotting is widely used to compare the relative levels of a target protein in a group of samples. Because this ratiometric analysis requires uniform loading of all samples, an internal loading control is detected and used for normalization.[11] For maximum accuracy, the target and internal control should be detected in the same context – at the same time, on the same blot, and in the same lanes.

4.4.1 Artifacts introduced by stripping and reprobing

When probing for multiple proteins using chemiluminescent detection, two options are available: stripping the blot and reprobing for a different protein, or simultaneously incubating the blot with all antibodies for multiplex detection. Although stripping and reprobing is very common, it's also time-

consuming and can introduce error. If stripping is incomplete, residual antibody will generate artifacts when the blot is reprobed; this is particularly troublesome if the two proteins migrate similarly on the gel. But complete removal of antibodies may be difficult to achieve. Stripping can damage or dissociate target proteins. It may also cause significant loss of sample proteins from the membrane – perhaps 25% or more of the target in a single stripping cycle.[2,25]

This issue was examined in a 2015 study, using immunoblots probed with phospho-ERK1/2 antibody and a GAPDH internal control.[1] The results showed that low-pH glycine stripping was very ineffective, with considerable residual antibody staining for p-ERK1/2 and GAPDH. Guanidinium stripping was more effective for removal of GADPH antibody, but a p-ERK1/2 artifact remained. An SDS stripping buffer with β-mercaptoethanol completely removed the GAPDH and p-ERK1/2 antibodies, but caused substantial loss of total protein from the membrane. This study describes stripping and reprobing of membranes as "a quantitative trade-off between antibody removal and total protein loss".[1] The potential effects of protein-specific factors (such as local overloading of the membrane, amino acid composition, and post-translational modification) on protein loss are an additional concern, although undocumented.

4.4.2 Multiplexing

Chemiluminescent detection is a single-color, single-channel method. If two proteins are detected simultaneously on a blot, chemiluminescent detection cannot identify or correct for antibody cross-reactivity artifacts. Proteins must be well separated on the blot, and controls should be used to check for antibody cross-reactivity. For multiplex detection, antibody incubations and detection are sometimes performed sequentially without stripping. However, this method is very time-consuming and may increase the membrane background.

Multiplex fluorescent immunoblotting makes normalization with an internal control simpler, more convenient, and more accurate. The blot is incubated with primary antibodies raised in different hosts. Secondary antibodies labeled with spectrally-distinct fluorescent dyes are then used to simultaneously detect the internal control on the same blot and in the same lanes as the target protein.[26-27] Blots are digitally imaged to detect the stable fluorescent signals (Fig. 3). Stripping is not required, co-migrating proteins can be used, and antibody cross-reactivity is easily identified. Stable expression of the internal control protein should be validated, as for any other fixed-point normalization strategy.

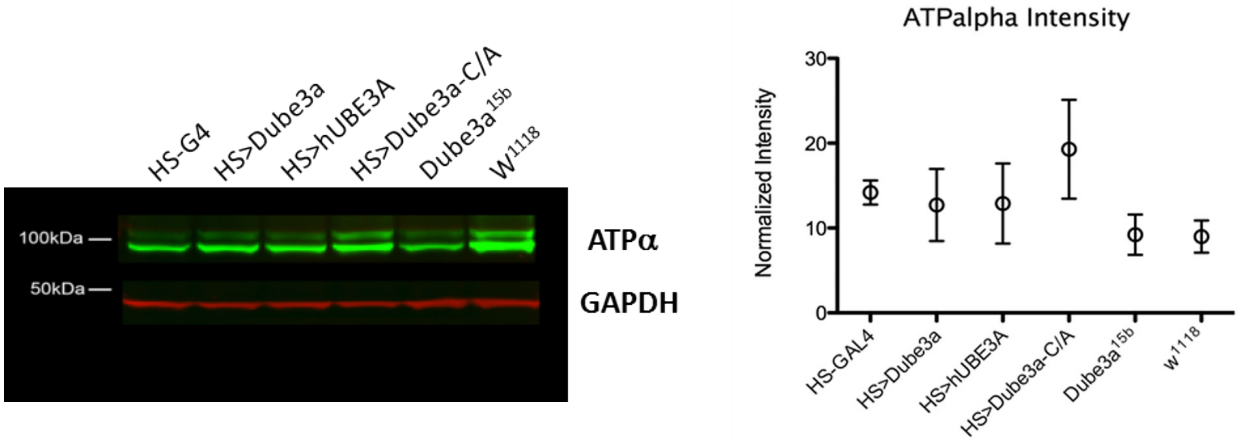

Figure 3. **Multiplex fluorescent Western blot detection with an internal control protein.** The stability and expression of ATPα (green) was examined in cell lines with differential expression of the ubiquitin ligase Dube3a. Normalization was performed with GAPDH as a loading control. NIR fluorescence imaging was used for detection. Normalized fluorescent intensities were graphed (n=3). Jensen et al. *PLoS ONE* 8(4): e61952.[28]

5. Saturation of highly abundant proteins

The impact of protein abundance and saturation on Western blot normalization is often overlooked. Saturation of strong internal control bands is especially problematic when a single internal loading control is used (normalization by fixed point). Many housekeeping proteins and structural proteins used as internal loading controls are highly abundant, but target proteins are often expressed at much lower levels.

Many factors contribute to saturation of strong bands, and saturation can occur at multiple points in the immunoblotting process. Saturation arises from limitations of protein transfer to the membrane, the detection chemistry used to generate signal, and the capacity of the imaging modality to record strong signals. When saturation occurs, strong signals will "plateau" and produce non-linear results that no longer reflect protein levels. These factors have a larger impact on data analysis when a single internal loading control is used for fixed-point normalization.

5.1 Membrane transfer

Overloading of samples is a common problem, but this parameter is not typically optimized. When an overloaded gel is transferred onto blotting membrane, highly abundant proteins can bind in layers on the membrane surface.[13,29] As a result, primary and secondary antibodies may only have access to the surface (top) layer of protein on the blot. This layering phenomenon contributes to underestimation of strong signals, and could interfere with detection of less abundant proteins. Even if other sources of saturation (such as detection chemistry and mode) are controlled, this layering effect is still able to cause saturation.

Membrane-related saturation is difficult to identify in Western blot images. A dilution series can quickly determine the upper limit, but the saturation point may be different for each target protein. Because it arises from the binding chemistry of proteins and blotting membranes, this type of saturation can occur with any detection chemistry or mode of imaging.

5.2 Detection chemistry

When internal control bands fall outside the linear range of detection, increases in protein level will not produce a proportional increase in signal intensity. Saturation artifacts can make strong bands appear to have similar intensity, hiding variations in protein levels and under-estimating the amount of protein present (see section 3.1, Fig. 2).

Large amounts of cell lysate are often loaded to facilitate detection of a target protein at endogenous levels, extremely overloading the abundant internal control protein (Fig. 4).[1-3,11,13] This pushes the internal control bands past the saturation point of the detection method. For accurate normalization, both the internal control protein and target protein must be detected within the linear range of the method used. When the abundance of the target and internal control are quite different, the linear range of detection and optimal sample loading will also be very different. This discrepancy makes quantitative analysis more complicated.

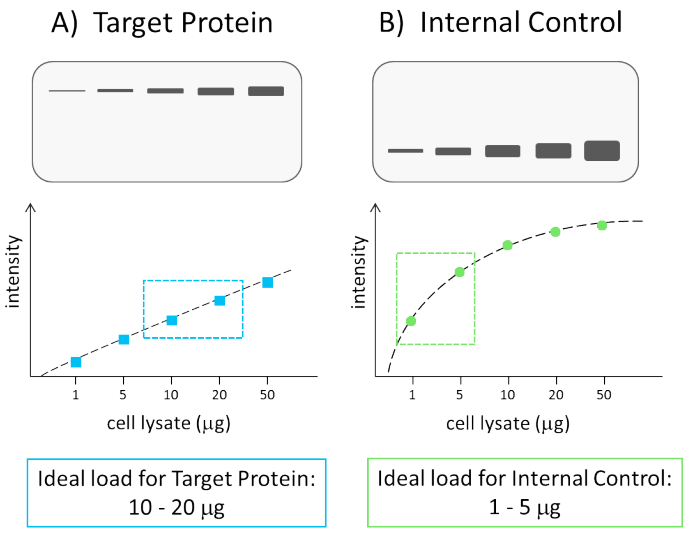

Figure 4. The linear range of detection may be different for the target protein and internal control protein. A) For the target protein depicted in this illlustration (blue), optimal sample loading would be ~10-20 µg lysate per lane. But for the internal control protein (B, green), that amount of lysate would be overloaded. Signals would be outside the linear range of detection. The amount of internal control protein would be underestimated, and could mask actual sample-to-sample variation in protein levels. Adapted from Taylor and Posch.[3]

The most commonly used detection chemistries for Western blotting are enhanced chemiluminescence (ECL) and NIR fluorescence. The reagents and detection chemistry used to visualize signals can be affected by saturation in several ways, and the enzymatic nature of chemiluminescent detection makes it particularly vulnerable to saturation artifacts.

5.2.1 Chemiluminescence

Chemiluminescent detection is an indirect, enzymatic method that typically uses secondary antibodies labeled with horseradish peroxidase (HRP) as an enzymatic reporter. This enzyme oxidizes a luminol-based substrate, transiently producing photons of light. Because signals are influenced by the kinetics of the enzymatic reaction, signal intensity does not always reflect the abundance of antigen.[1,24,30]

This detection chemistry provides high sensitivity, but may exaggerate or underestimate signal intensity. The enzymatic reaction is not necessarily proportional or linear, and a two-fold change in protein abundance may be measured as an exaggerated five-fold or greater increase in signal intensity.[1] But enzyme/substrate dynamics also cause underestimation of stronger signals, as the "runaway reaction" drives them to saturation.[1-3,11,13] These contradictory effects occur because the chemiluminescent reaction amplifies signal by turnover of substrate at the site of antibody binding. Signal amplification creates significant problems for quantitative analysis.

Local concentrations of enzyme and substrate can greatly impact the rate of the chemiluminescent reaction. The local concentration of substrate varies continuously throughout imaging, as the enzymatic reaction proceeds and consumes substrate.[25,31-32] Thus, substrate availability can vary widely across the membrane and is constantly in flux. When antibodies bind to a low-abundance protein on the blot, the local concentration of HRP enzyme is low and an excess of substrate is maintained. But an abundant protein such as a loading control will bind more antibody molecules, creating a high local concentration of HRP that consumes the reagent at a higher rate. Although excess substrate may initially be present, the kinetics and reaction rate will change as substrate is rapidly consumed.[31]

High local concentrations of enzyme-labeled antibody complex can also lead to loss of signal in several other ways. Side products of the peroxidase-catalyzed reaction may become oxidized and precipitate, causing the membrane to "brown out" and begin absorbing emitted light.[1] Additionally, the HRP enzyme itself can be inactivated by prolonged exposure to substrate, as free radicals produced by the reaction bind to the enzyme and reduce its ability to interact with substrate.[2]

The enzyme/substrate dynamics of chemiluminescent detection introduce variability and often generate non-linear responses. The limitations of this detection chemistry make quantitative analysis more challenging, and may affect the reproducibility of Western blot results – particularly when a single internal control is used for normalization by fixed point. These limitations apply to chemiluminescent Western blots in general, regardless of the method used to document the signals.

5.2.2 Fluorescence

Fluorescent detection is generally considered the most accurate method for quantitative immunoblotting.[26-27,33-34] Fluorescent immunoblotting is a direct, non-enzymatic method that uses secondary antibodies covalently labeled with fluorescent dyes (fluorophores), typically in the near-infrared (NIR) spectrum. Fluorophores are retained at the site of antibody binding, and generate signals across a wide linear range when exposed to excitation light of appropriate wavelength. When the appropriate fluorescent dyes are selected, fluorescent signals are very stable; blots can be stored for days or months and re-imaged later.[27]

Fluorescent detection is inherently more reproducible than enzymatic methods, because signal intensity is not affected by timing or enzyme/substrate dynamics.[1,26,35] In heavily loaded samples, tightly packed fluorophore-labeled antibodies do have the potential for self-quenching.

5.3 Imaging modalities

The detection technology that records your results may impose its own limitations. Signal saturation is extremely common with some detection modes; it truncates the linear and dynamic ranges and may greatly affect the accuracy of normalization. Wide linear range is an important consideration for quantitative immunoblotting – and is absolutely critical for comparison of samples when target protein abundance is variable.[1,33]

Chemiluminescent Western blots are documented by exposure to x-ray film or with a CCD imaging system. Both modalities are affected by saturation, but with very different mechanisms. Digital imaging is used to document fluorescent immunoblots using CCD, PMT, or APD-based imaging systems.

5.3.1 Film exposure of chemiluminescent blots

Although film is commonly considered to provide the highest quality Western blot image, this is a misconception.[13] In fact, film-based detection has two fundamental limitations that affect the analysis and reproducibility of immunoblotting data. It provides an extremely narrow linear range of detection, roughly 4-10 fold (Fig. 5A); and rapid saturation of strong signals makes it very difficult to accurately determine the limits of detection, particularly the upper limit.[13] Film can exaggerate small differences in abundance, while also masking sample-to-sample variability of strong bands.[1]

Plateau and saturation of strong signals are artifacts of the photographic emulsion that coats the film. Photons of light from the chemiluminescent reaction activate individual silver grains in the emulsion, which are then converted to black metallic silver to create a visible image. A strong signal

will quickly activate the majority of silver grains in that area of the film. Film responsiveness slows and eventually saturates, as new photons of light become less likely (and eventually unable) to activate additional silver grains in that local area. This effect, known as high-intensity reciprocity failure, causes the rapid plateau and saturation of strong signals commonly observed in film exposures.[36-38] As a result, the intensity of strong bands may be significantly underestimated and actual changes in relative protein level will not be detected.

The plateau and saturation of strong signals are particularly risky because they are largely undetectable to the user. Without a dilution series of controls on every blot, it's extremely difficult to determine the upper limit of detection. Film response becomes compromised before it reaches the point of saturation – but there's no reliable way to determine when this begins to affect data output. Because of these limitations, film should not be used for quantitative immunoblotting.[1,3,13,33]

5.3.2 CCD imaging of chemiluminescent blots

CCD imaging typically offers a wider linear range of detection than film (Fig. 5B).[3,10,13,26] Many CCD systems are able to detect some lower-intensity signals without extensive saturation of strong signals. However, saturation does occur with some imagers, and the resulting linear range may be less than 100-fold.[3,10,26] Sensitivity and linear range will vary, depending on the imaging hardware.

Saturation of CCD images is often represented by the imaging software as colored pixels. CCD imaging provides improved image quality and clarity, and eliminates the error and variability introduced by scanning of exposed films for densitometry.[39] Overall, digital imaging more accurately represents band intensity and protein abundance than film exposure (Fig. 5B). If immunoblot accuracy is optimized and results are carefully interpreted, this method can be useful for quantitative immunoblot analysis.[1-3] Even with a digital imaging modality, Western blot signals are still subject to the nonlinear enzymatic effects of chemiluminescent detection that may affect outcome.

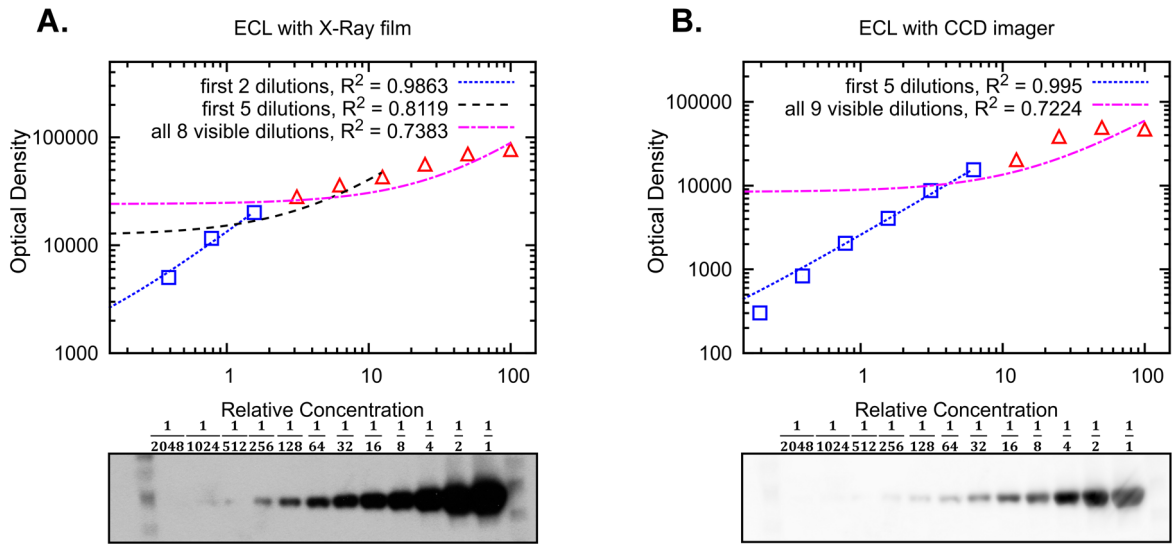

Figure 5. Film imaging provides a much narrower linear range than digital imaging with a CCD system. A) Two-fold serial dilutions were detected on replicate Western blots, using ECL substrate. A) Film exposure. Only the first 2-3 dilutions (blue squares) fell within the linear range, with $R^2 = 0.9863$. All other dilutions were subject to saturation (red triangles). B) CCD imager (longest exposure with no saturation observed). The linear range spanned 5 dilutions (blue; $R^2 = 0.995$). Image clarity was improved; stronger bands were more distinct and did not overlap. Membrane overloading and detection chemistry effects are likely both contributing to the saturation observed in both A and B. Degasperi et al. *PLoS ONE* (2014) doi:10.1371/journal.pone.0087293.[10]

5.3.3 Digital imaging of fluorescent blots

Fluorescent immunoblotting is best performed with near-infrared (NIR) fluorescent dyes and imaging systems.[1,5,12,26,30,33-35] Background autofluorescence of transfer membranes and biological samples is low in the NIR spectral region, enabling high sensitivity. In contrast, high background fluorescence in the visible spectrum causes poor signal-to-noise ratios, limiting sensitivity and performance. NIR fluorescent signals are typically documented by measuring fluorescence intensity with a laser-based APD scanner or CCD imager. The sensitivity and linear range of this method are also influenced by the imaging instrumentation used. The excitation light source and optical system affect sensitivity and background levels. The ability to detect low-intensity signals without saturation of strong signals, which is crucial for a wide linear dynamic range, is highly dependent on the imaging hardware and varies from 2.5 to > 6 orders of magnitude (Fig. 6).

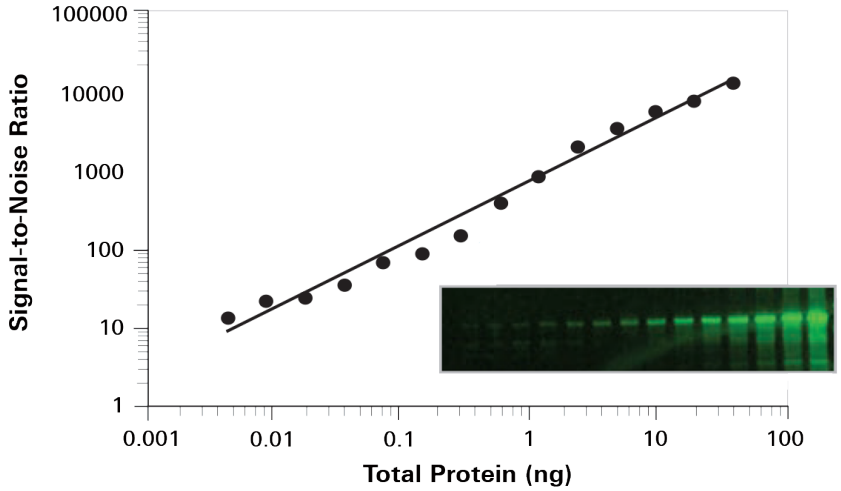

Figure 6. NIR fluorescence imaging provides a very wide linear dynamic range of detection. Two-fold serial dilutions of purified transferrin were detected by NIR fluorescent immunoblotting, using a laser APD scanner. All 14 dilutions, from 40 ng to 4.8 pg, fell within the linear range without signal saturation (an 8000-fold range).

6. Using internal loading control proteins for immunoblot normalization

The most commonly used internal loading control strategy is detection of a second, unrelated protein on the same blot (normalization by fixed point; described in section 3.1). An ideal internal control protein would be stably expressed in all cell types, with its expression level unaffected by changing experimental conditions. Housekeeping proteins such as actin, tubulin, and GAPDH are widely used and published as internal control proteins, but their expression may not be as stable as previously hypothesized.[4-9]

Multiplex immunoblot analysis of post-translational modification is a different type of internal control. The blot is incubated with a mixture of modification-specific and pan-specific primary antibodies against the target protein, and fluorescently-labeled secondary antibodies are used to detect and discriminate the two signals.[26] This ratiometric analysis approach uses the target protein as its own internal loading control (discussed in section 6.3).

6.1 Housekeeping proteins as internal loading controls

Use of housekeeping proteins as loading controls is based on the concept of "housekeeping genes," stable endogenous controls used for data normalization in gene expression assays that measure transcript abundance. A housekeeping gene is defined by its constitutive expression throughout all cells of an organism. Ideally, such a gene would not be subject to regulation and would be expressed at a constant level in every cell type.[40]

Published observations of relatively stable gene expression of housekeeping genes such as actin, tubulin, and GAPDH across cell types and tissues have led to widespread use of their protein products as internal loading controls for quantitative immunoblotting. But in a cellular context, expression of these proteins can be much more variable than expected.[4-9]

Validation of the internal reference protein is very important, to confirm stable expression across the relevant cell types, tissue types, disease states, and/or experimental treatments.[1-2,4-9,19] A housekeeping protein such as actin or tubulin may not be the most appropriate choice for every experiment, and fixed-point normalization with a single control protein warrants consideration of additional factors (as discussed in section 3.1).

Biological and methodological factors contribute to the variability of housekeeping proteins as internal controls for immunoblot analysis:

- Gene expression does not reliably predict the abundance of the corresponding proteins[41-42] (see section 6.1.1).
- Housekeeping protein expression is not always constant across cell types and tissues. It may be affected by a variety of biological factors, including tissue type, growth conditions, stage of development, and disease (see section 6.1.2).
- Housekeeping proteins are typically very abundant. The resulting strong bands frequently cause signal saturation, which reduces the accuracy of detection and under-estimates the actual amount of internal control protein in each lane (as discussed in section 5).

6.1.1 Gene expression levels do not reliably predict protein abundance

For many proteins, mRNA levels do not reliably predict protein abundance.[43-47] A 2014 mass spectrometry study built a draft of the human proteome.[47] Published mRNA abundance values were compared to quantitative protein-expression profiles for twelve human tissues. Correlation between mRNA and protein abundance was weak in all tissues, with an average RS value of 0.41 +/- 0.07 (Table I).

Tissue	R_s
Uterus	0.34
Kidney	0.56
Testis	0.33
Pancreas	0.38
Stomach	0.45
Prostate	0.41
Ovary	0.39
Thyroid	0.31
Adrenal gland	0.47
Salivary gland	0.46
Spleen	0.39
Esophagus	0.45
Average R_s	**0.41 +/- 0.07**

Table 1. Spearman correlation coefficients for mRNA versus protein abundance plots for proteins in 12 human tissues. mRNA values were drawn from an existing database, and were not measured in the same tissue samples used for protein analysis. Wilhelm et al.[47]

Because post-transcriptional regulatory mechanisms strongly influence protein abundance, gene expression observations should be validated at the protein level to confirm stable expression of any housekeeping protein used as an internal loading control.

6.1.2 Expression of housekeeping proteins may be influenced by biological factors

Several recent studies have raised concerns about the use of common reference proteins as loading controls, reporting differential expression of housekeeping proteins in certain circumstances.[2,4-9,11] One example is shown here. Examination of human adipose tissue samples demonstrates highly variable expression of beta-tubulin in the same tissue type from obese subjects and non-obese controls (Fig. 7).[7] Subject-to-subject variation was observed within each group, and also between the two metabolic states.

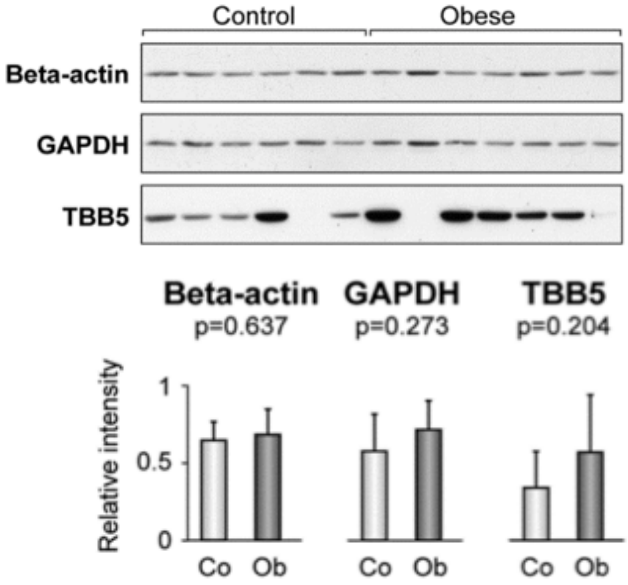

Figure 7. Variable expression of housekeeping proteins is observed in human adipose tissue from non-obese (control) and obese subjects. Western blot analysis of beta-actin, GAPDH, and tubulin beta chain (TBB5) expression in omental fat samples from 6 non-obese controls (Co) and 7 obese (Ob) subjects. Relative band intensity is expressed as mean ± SD. Pérez-Pérez et al. doi:10.1371/journal.pone.0030326.[7]

Developmental status is another important source of variable protein expression. To identify an appropriate housekeeping gene for use as an internal control in a retinal development study, expression profiles of three internal control candidates were examined across distinct developmental stages (Fig. 8).[6] From embryonic to post-natal development, each housekeeping protein displays a unique expression profile. Alpha-tubulin levels rise during early post-natal development, but then fall. Beta-actin is highly expressed in embryonic development, but expression drops off dramatically in post-natal stages. MAPK1 was also examined, and unlike the traditional housekeeping proteins, expression of this signaling protein was relatively constant throughout these stages of retinal development. The authors conclude that of the proteins tested, MAPK1 would be the most suitable internal control protein for retinal development.

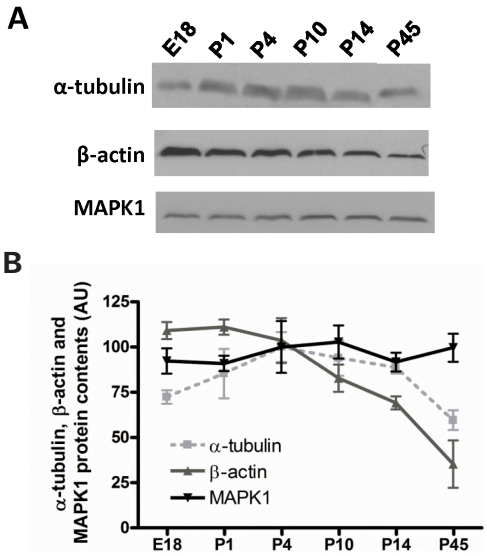

Figure 8. Variability of housekeeping protein levels throughout retinal development in rat. A) Western blot analysis of α-tubulin, β-actin and MAPK1 is shown for distinct retinal developmental stages, from embryonic to post-natal (E18, P1, P4, P10, P14, P45). B) Densitometric analysis of β-actin, α-tubulin and MAPK1. Values expressed as Mean ± SEM for 3 experiments. β-actin displayed the greatest variability during development, with CV = 36.6%. Rocha-Martins et al. *PLoS ONE* 7(8): e43028 (2012). doi:10.1371/journal.pone.0043028.[6]

Changes in protein expression also arise from environmental factors. Diurnal rhythm is one example. The neurexins are neuron-specific presynaptic proteins that participate in regulation of the excitatory/inhibitory balance in synapses. A 2012 study explored the hypothesis that neurexin expression might exhibit circadian patterns connected to the light:dark cycle.[48] Diurnal variations were demonstrated for expression of neurexins, exon splicing, and postsynaptic scaffolding proteins. Figure 9 shows one example.

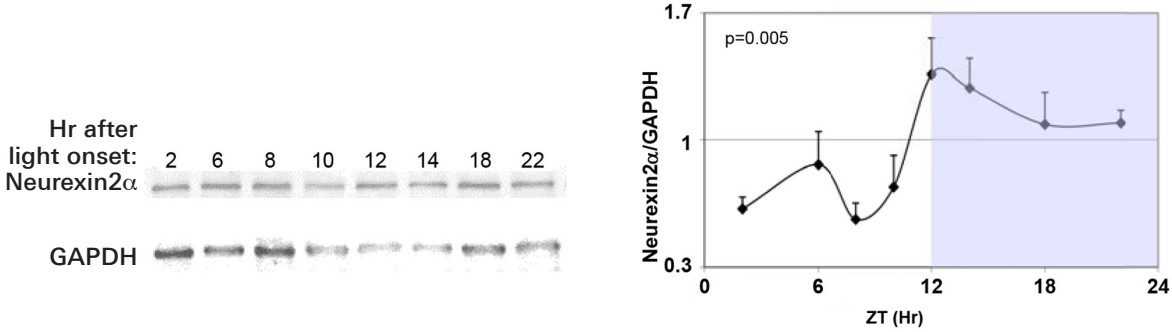

Figure 9. Diurnal rhythm influences neurexin 2α protein expression in mouse neurons C3H/J mice kept at 12:12 hours light:dark schedules were sacrificed at different times after light onset (0 hr). Expression of neurexin 2α (a synaptic receptor protein of the vertebrate nervous system) was examined by fluorescent immunoblotting. A) Representative blot shows levels at various times after lights-on. B) Values expressed as Mean +SEM relative to GAPDH (N = 3 per time point). Purple shading indicates the dark phase. Shapiro-Reznik eta l. *PLoS ONE* 7(5): e37894 (2012). doi:10.1371/journal.pone.0037894.[48]

Expression of reference proteins may even vary in different portions of a single tissue. The levels of reference proteins were examined in proximal and distal portions of sciatic nerve from the same mouse.[5] Although total protein concentration was constant, relative expression of actin and neurofilament-L (NF-L) was different in proximal and distal tissue samples. Abundance of actin was ~50% lower in the distal sample, whereas NF-L abundance was nearly 8-fold higher in the distal sample than the corresponding proximal sample. Although this study was performed with a structurally asymmetrical tissue, it emphasizes the importance of reference protein validation and implicates tissue sampling procedures as a possible source of variability in quantitative immunoblotting.

6.2 Saturation of strong bands

Housekeeping proteins are typically expressed at high levels. The resulting strong bands often cause saturation of signal, particularly when chemiluminescent detection is used.[2-3,11,13] Quantification of saturated bands generally underestimates the amount of internal reference protein in each sample. This effect may hide actual sample-to-sample variation in protein abundance,[1,11] as discussed in section 5.

6.3 Multiplex analysis of protein modifications: a different kind of internal control protein

Multiplex fluorescent detection is another option for immunoblot normalization. This approach is particularly useful for relative analysis of post-translational modifications such as phosphorylation. The method combines two primary antibodies raised in different hosts: a phospho-specific antibody (or other modification-specific antibody) and a pan-specific antibody that recognizes the target protein regardless of its modification state.[1,34,49-51] Fluorescently-labeled secondary antibodies are used to simultaneously detect and discriminate the two signals with digital imaging.[26] Phospho-signal is then normalized against the total level of target protein, using the target protein as its own internal control.

This method is widely used, and published examples are shown in Figures 10 and 11.[49-50] In validation experiments, Bakkenist et al. examined the possibility of binding interference from combined phospho-specific and pan antibodies, but detected little or no effect.[34] An additional, independent loading control may be used to verify sample loading, if desired.

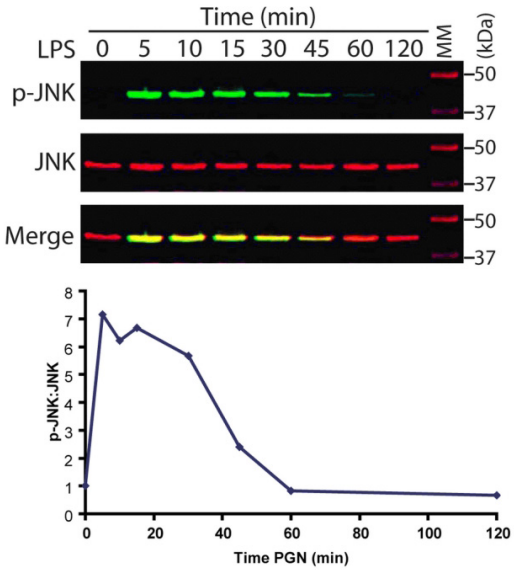

Figure 10. Time course of JNK phosphorylation in response to treatment with LPS. Multiplex fluorescence was used for ratiometric analysis of phosphorylation. Phospho-JNK is shown in green (800 nm) and total JNK in red (700 nm). Yellow indicates overlapping signals in the merged image. Reprinted with permission from Bond et al. *Biol Proced Online*. 10(1): 20-28 (2008).[49]

Two-color phospho-analysis has several advantages. Sample volume is reduced, and analysis is streamlined because all data are derived from a single blot without stripping and reprobing. Analytical confidence is increased by detection of unmodified and modified forms of the target protein on the same blot, in the same lane. No error is introduced by stripping and reprobing. Ratiometric analysis improves accuracy by correcting for loading and sampling error, and eliminates the uncertainty and variability that housekeeping proteins can introduce. Figure 11 illustrates the importance of detecting the internal control protein in the same lane of the same blot.[50] The Akt band in the last lane (far right) is irregular, likely because of a small bubble present during membrane transfer (Fig. 11, arrowhead). With multiplex detection, the same bubble is observed in the p-Akt (green) and total Akt (red) images and this transfer artifact is easily corrected.

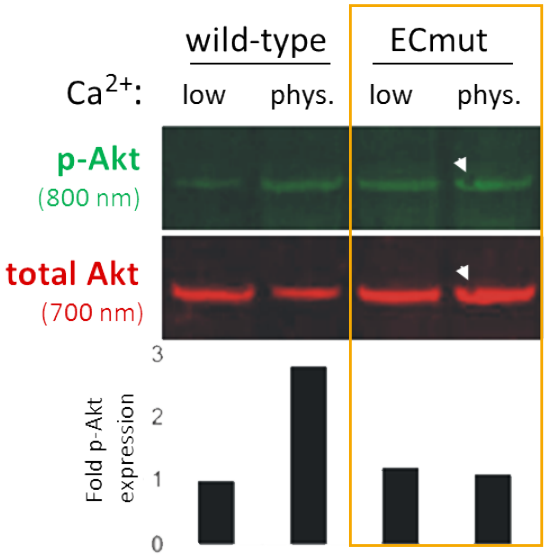

Figure 11. Loss of E-cadherin function affects PI3K/Akt signaling. Human urothelial cells with a dominant negative mutation in E-cadherin (ECmut) were grown in low or physiological Ca^{2+} conditions. Lysates were examined by multiplex Western blotting. Physiological Ca^{2+} increased phospho-Akt levels in wild-type cells but not ECmut cells (box). Arrowheads indicate a transfer artifact. Georgopoulos et al. *PLoS ONE* 5(10): e13621 (2010). doi:10.1371/journal.pone.0013621.[50]

Although ratiometric analysis is used to monitor and compare the relative phosphorylation response in a group of samples, relative comparisons do not indicate phosphorylation stoichiometry. Calculation of stoichiometry requires electrophoretic separation of phosphorylated and total protein forms; it cannot be determined using antibodies with different affinity characteristics.[1]

Analysis of other post-translational modifications, including ubiquitination, glycosylation, and acetylation, can also be performed in this manner.[28,52-53]

Chemiluminescent detection can make modification-specific immunoblotting much more difficult. Blots are sometimes stripped and reprobed with a pan-specific antibody, but it's essential to confirm that the original modification-specific antibody is completely removed. Any residual antibody will create artifacts when the blot is reprobed. Complete removal of antibodies can be difficult, and may require harsh stripping conditions that cause substantial loss of transferred proteins from the membrane.[1-2,25] Methods such as two-color fluorescence or detection of replicate blots are alternatives to stripping and reprobing.[1]

6.4 Internal loading control proteins: strengths, limitations, and considerations

For accurate normalization by fixed point, an internal control protein should meet several requirements:

Its abundance must be relatively constant for the conditions and samples relevant to the experiment (cell and tissue type, drug treatments, growth conditions, disease state, etc). A housekeeping protein may meet this need, or an alternative internal reference protein may be selected. Validation is essential to confirm stable protein expression in the samples, in your lab, and with your antibodies.[2] If the internal control is highly abundant, sample loading should be optimized to ensure that the protein is detected within the linear range and signals are not saturated. Target protein and internal control protein signals must both fall within the linear range of detection.

Signal intensity for the reference protein must display a linear relationship with the abundance of the reference protein. The upper and lower limits of the linear range can be determined empirically, and will be different for each protein detected.[1-3] At the upper end of the range, signals will become saturated and variations in protein abundance will not be detected. At the lower end of the linear range, low-intensity data points are affected by random noise and are generally not appropriate for normalization. Optimization of sample loading can help you identify and maximize the linear range of detection.

If a single internal reference protein does not meet these requirements, a multi-protein strategy (normalization by sum; discussed in section 3.2) may be needed.[1,11] Stripping and reprobing of blots can introduce detection artifacts and cause loss of blotted proteins from the membrane; it should be used with caution.[1-2,25]

7. Using total protein stains for immunoblot normalization

Concerns about the reliability of internal loading control proteins have prompted interest in alternative approaches for normalization. Total protein normalization, using a protein stain to detect the total protein in each lane of the gel or blot, is an increasingly popular option.[5,9,11,20-23,54-57] For each lane, the signal intensity of all proteins in the lane is measured in aggregate; that value is used in normalization calculations to represent the total protein content of the sample.

7.1 Total protein staining

This emerging method eliminates many of the challenges associated with internal control proteins. Total protein staining is a direct measure of the total amount of sample protein in each lane, and does not rely on an internal reference protein as an accurate indicator of sample concentration. Taking a more direct approach to assessment of sample proteins may increase the accuracy of normalization. And unlike an internal control protein, total protein staining does not require validation for each experimental context, biological system, and primary antibody. A total protein stain should produce a linear increase in signal intensity in response to increasing protein concentration. It should also correct for variation at all points in the Western blot process, including gel loading and transfer to membrane. And it must be compatible with subsequent immunodetection of the blot.[11,20-22]

Some protocols involve staining of a gel for comparison of total protein.[5] This is an excellent method for standardization of sample loading, but does not address potential inconsistencies in membrane transfer. Other protocols call for direct staining of the membrane after transfer.[20-23,55,57] This provides a more

detailed assessment of total lane density on the blot by accounting for variability in transfer efficiency, protein binding, and other factors. However, the membrane stain must be compatible with subsequent immunostaining. In some protocols, the membrane is stained after immunodetection is complete.[10] Alternative approaches have also been proposed; these "stain free" methods involve labeling of total sample proteins with a detection moiety.[13,23,56] Protein labeling occurs prior to or during gel electrophoresis (see section 7.4).

7.2 Coomassie staining of total protein for normalization

Although Coomassie staining has traditionally been quantified by densitometry, Coomassie Blue is also an excellent near-infrared (NIR) fluorophore. Gel imaging with an NIR fluorescence laser scanner generates strong 700 nm signals, with a wide linear response to protein concentration. Fluorescent signal is proportional to protein content across a range of 10 ng to 20 µg per band.[54] Because fluorescence is induced by binding to protein, background is very low and sensitivity is excellent.

Replicate gels or blots can be stained with Coomassie to standardize sample loading. It can also be used to stain the Western blot membrane after immunodetection.

7.2.1 Coomassie staining of replicate gels

The replicate gel approach was validated in a 2014 study.[5] Wishart and colleagues explored Coomassie staining as an alternative to commonly-used internal control proteins. First, they evaluated the linear range and sensitivity of total protein staining (Fig. 12). Coomassie staining (detected by NIR fluorescence and pseudo-colored in red) produced a total protein signal that was linear across the entire range of protein concentrations (Fig. 12, red line). Coomassie results directly correlated with protein concentration data from BCA assays (Fig. 12, blue line). Total protein analysis by Coomassie staining provided a valid, linear readout of sample concentration.

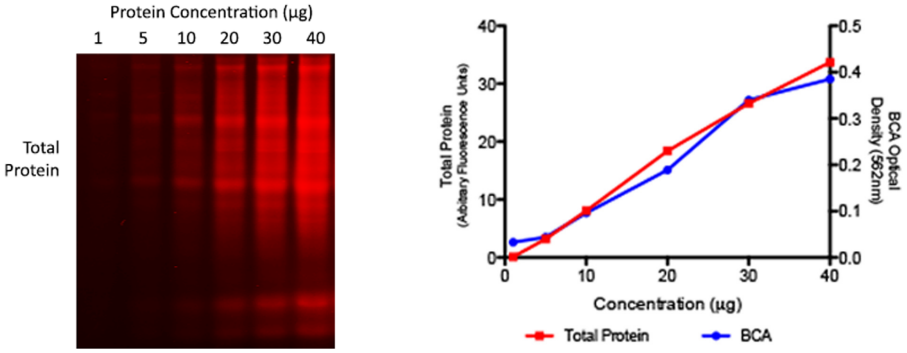

Figure 12. Coomassie staining of total protein is an accurate measure of protein load. Total protein was detected in whole brain homogenates with Coomassie stain and NIR fluorescence imaging. Coomassie signals (R^2=0.996) were directly correlated with BCA assay results (R^2=0.979) across a wide range of protein concentration. Eaton et al. doi:10.1371. journal.pone.0072457.[5]

Beta-actin expression was then examined in various tissues from wild-type mice, using total protein analysis for normalization.[5] Actin was detected by Western blotting in samples from six different tissues and was highly variable (Fig. 13A, green bands). In the quantitative analysis, Coomassie staining showed that total protein levels were relatively consistent but actin expression varied widely (Fig. 13B). Total protein is a more appropriate loading control than actin for cross-tissue comparison of these mouse samples.

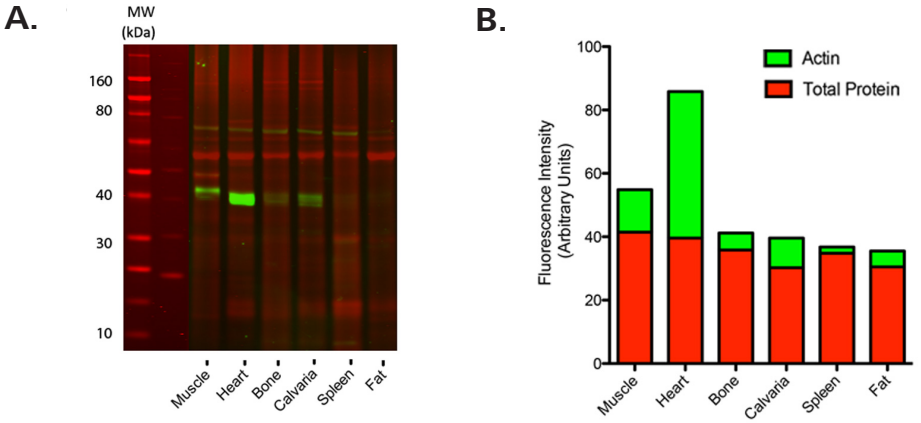

Figure 13. Expression of β-actin is highly variable in different mouse tissues. A) Actin expression was assessed by immunoblotting in tissue samples from muscle, heart, bone (femur), calvaria, spleen, and fat (gonadal). Multiplex immunoblot analysis with NIR fluorescence demonstrates variability of actin expression (green) in these tissues. A total protein gel image (red; overlaid on blot) was used to confirm the accuracy of protein loading across the samples. B) Stacked bar graph shows the comparative variability of β-actin (green bars) and total protein stain (red bars) for each tissue. Eaton et al. doi:10.1371/journal.pone.0072457.[5]

7.2.2 Coomassie staining of immunoblots

The replicate gel strategy is simple and useful, but does not account for variability in membrane transfer. Welinder and Ekblad demonstrated conventional Coomassie staining of PVDF blots after immunodetection as an internal control for protein loading.[55] The linearity of this method was superior to antibody-based detection of GAPDH, and more sensitive than the reversible total protein stain, Ponceau S. Although the blot had been blocked with nonfat dry milk, blocking proteins caused only negligible background staining and protein bands were clearly resolved. This post-staining approach corrects for variation in membrane transfer, and guarantees that total protein staining will not affect antibody binding and detection.

7.3 Other options for total protein staining of membranes

A variety of stains can be used to detect total protein on blotted membranes.[1-3,5,54-57] Some stains are reversible and can be removed from the blot prior to immunodetection. Depending on the binding chemistry, reversal of staining may require changes in pH or solvent hydrophobicity.[22] Commercially available total protein stains with signals of various colors and wavelengths are available. Depending on the stain, detection and quantification may be performed by densitometry or fluorescence imaging.

Several factors should be considered when you choose a total protein stain for normalization by sum. The sensitivity and linear range of each stain are different, and these parameters may affect your experiment. Although some blot stains are quite sensitive and able to detect low-nanogram amounts of protein, others are orders of magnitude less sensitive. Linear response range should be empirically tested to make sure it is appropriate for your sample concentrations.[2] Certain stains are compatible with only one type of membrane, nitrocellulose or PVDF. Some stains may cause increased membrane background,[1] and significant residual staining of protein bands is sometimes observed. The additional time required to stain and destain the blot is another consideration. It's also possible that use of a reversible stain could somehow affect certain epitopes. This is also a concern with staining methods that chemically modify sample proteins by covalent attachment of a detection molecule (section 7.4).[56] Coomassie staining performed after immunodetection would eliminate this possibility,[55] as discussed in section 7.2.2.

7.4 Alternative methods for total protein normalization

7.4.1 Stain-Free™ technology

The "Stain-Free" system from Bio-Rad takes a different approach to total protein detection – chemical labeling of the sample proteins.[3,13,56] The labeling moiety, a trihalo compound, is incorporated into the proprietary pre-cast gel. When this gel is exposed to UV light, the trihalo label covalently binds tryptophan residues in the sample proteins and forms a cross-linked fluorescent product. Fluorescently-modified proteins are then detected with a CCD imager and UV illumination.

Total protein can be assessed in the gel with good sensitivity across a linear range of 1-35 µg of cell lysate (Fig. 14A).[13] However, sensitivity is greatly reduced and the linear range is much narrower for detection of proteins on the transferred membrane. Loss of sensitivity is caused by high membrane background, due to autofluorescence of the PVDF membrane at UV wavelengths.[13] Membrane imaging of Stain-Free labeled proteins offers a linear range of only 10-70 µg (Fig. 14B), similar to Ponceau S.

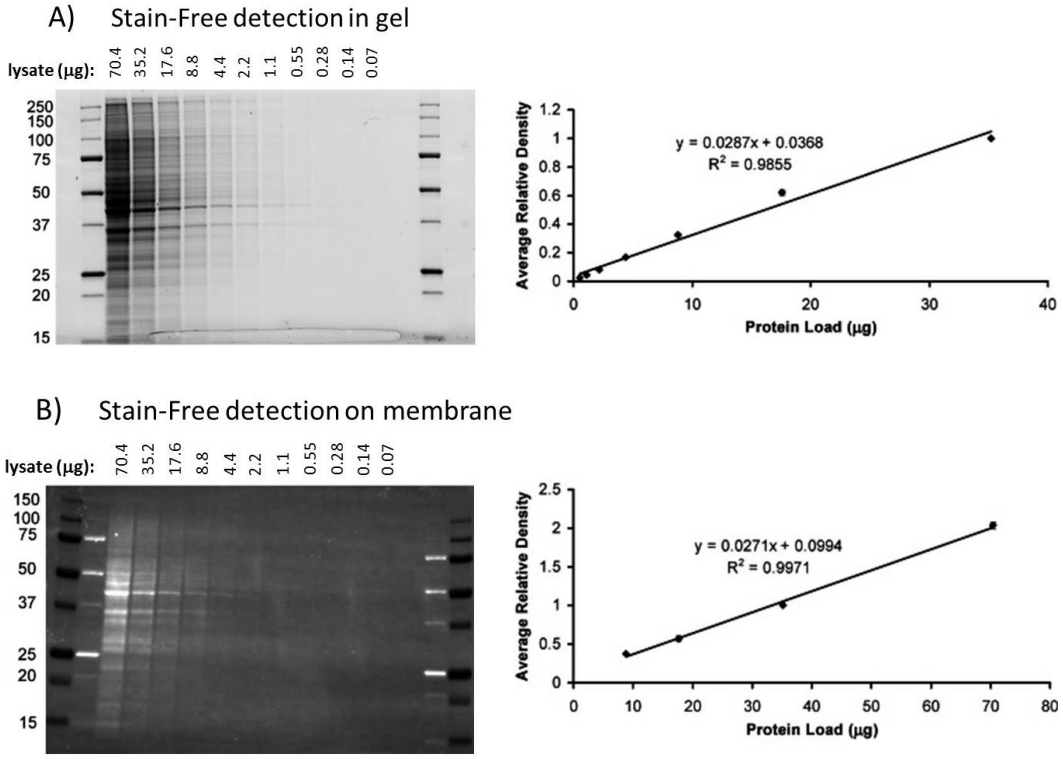

Figure 14. Stain-Free imaging of transferred membranes has a much narrower linear range than imaging of gels. Two-fold serial dilutions of HeLa cell lysate were loaded. A) Detection of Stain-Free labeled proteins in the gel. Graph displays the average relative lane density of the total protein load for three gels. MW markers were run in the first and last two lanes. B) Detection of Stain-Free labeled proteins after membrane transfer. Membrane autofluorescence is high at UV wavelengths, limiting the detection sensitivity and linear range. Taylor SC et al. 2013. *Mol Biotechnol.* 55:217-26.[13]

7.4.2 Limitations of the Stain-Free™ technology

The Stain-Free technique is convenient but has notable limitations. Unlike conventional stains that can be washed away, the Stain-Free chemistry is covalently, irreversibly bound to your protein samples. Chemical modification of tryptophan residues may interfere with immunodetection of some epitopes or have other unintended downstream effects.[3,13,56]

Although Stain-Free chemistry performs well for gel imaging, optimal normalization also requires correction of variability introduced by membrane transfer. Inconsistent transfer efficiency can result in two- to four-fold changes in signal across the blot.[10,13] If uncorrected, these effects add uncertainty to the data and may increase the coefficient of variation (CV).

The limited sensitivity of Stain-Free chemistry for total protein detection on membranes may also limit its usefulness. The published limit of detection is ~10 µg/well (Fig. 14B).[13] Given that sample loading of 10-25 µg/well is common,[1-2,13] total protein levels may routinely fall in the lower end of the Stain-Free linear range, near the lower limit of detection. These low-intensity data points may be unsuitable for normalization. Low-intensity normalization signals have larger CVs and may introduce uncertainty, increasing the mean CV of the normalized data and producing false negatives (undetectable differences between protein samples).[10-11] This may be a concern in the context of normalization by sum, because the redistribution of uncertainty may further increase the CV of low-intensity data points.

Increasing the amount of sample protein loaded on the gel will not overcome this sensitivity limitation. As discussed previously, overloaded samples may produce strong Western blot signals that become saturated. High-intensity data points outside the linear range will also affect the accuracy of normalized data.[10] It may be possible to boost the Stain-Free signal intensity by extending the UV crosslinking step, but this increases the extent of covalent modification and the risk of interference with antibody binding.[56]

This method also offers less flexibility than conventional stains. Proprietary gels and specific equipment for transfer and imaging are needed, requiring you to change established lab protocols.[21] The predetermined workflow makes it difficult to adjust the staining protocol to fit the sensitivity and linear range needs of different experiments or target proteins.

7.4.3 Total labeling of protein samples with reactive dye

Pre-labeling of protein samples with a lysine-reactive fluorescent dye is another option for normalization. In this method, amine-reactive fluorescent cyanine dye is mixed with the total protein sample prior to electrophoresis.[22] Reactive dye covalently binds to lysine residues in the sample proteins, forming a stable conjugate. The extent of protein labeling is controlled by the reaction stoichiometry, using a small amount of amine-reactive dye to label approximately 5% of total protein.

As with the Stain-Free method, covalent modification of sample proteins could affect downstream analysis. Conjugation of the cyanine dye to sample proteins may alter epitopes and affect antibody binding. Partial labeling of a complex mixture with reactive dye is not straightforward, and the extent of protein labeling in this method is inherently dependent on amino acid composition.

7.5 Total protein normalization: strengths, limitations, and considerations

An ideal total protein stain for normalization by sum has several key characteristics:

Linear signal output in response to sample concentration. Many protein stains can meet this need, but linear range and sensitivity can vary widely and may be specific to the type of membrane used. These parameters should be considered when choosing a stain for total protein normalization.

Corrects for variation at all points in the Western blot process. Gel-staining approaches are useful and inexpensive, but do not correct for membrane transfer variability. Staining of total protein on the blotted membrane may be the most accurate approach for Western blot normalization. Several good staining options are available, but the "best" choice for membrane staining is unclear. Sensitivity and linear range should not be sacrificed to enable detection on blotted membranes.

Compatible with immunodetection of target proteins. Membrane stains are a promising option, but staining of replicate gels or staining of blots after immunodetection may also meet researchers' needs. For a number of membrane stains, no downstream effect on immunodetection has been demonstrated. Methods that covalently modify target proteins for fluorescent detection have the potential for downstream interference.

Total protein staining after membrane transfer is emerging as possibly the most reliable and accurate method for normalization of Western blot data. This normalization strategy will require staining reagents that are affordable and flexible, offer a wide linear range of detection and excellent sensitivity on membranes, and add minimal handling time and effort to the quantitative immunoblot process.

8. Conclusions and future directions

Internal loading controls and normalization are critical for quantitative immunoblotting. An accurate loading control will display a linear relationship between signal intensity and sample concentration. When implemented, an effective normalization strategy should correct for variability in all stages of the immunoblotting process, including the transfer of sample proteins to membrane. It should also be compatible with immunodetection of target proteins and other types of downstream analysis.

As researchers detect and interpret subtle changes in protein samples, accurate normalization is becoming increasingly important. Although use of an internal control protein such as actin or tubulin is perhaps the most common strategy, recent studies are raising questions about this approach. The evidence indicates that an internal reference protein should be validated in the relevant experimental context, to confirm stable expression. But these proteins are still widely used for normalization, and routinely published without validation data. Many researchers may be unaware that an internal reference protein is not always sufficient for normalization, or that validation of the control protein is important. We see a constant stream of new publications that use this familiar, widely-accepted internal reference protein approach for quantitative Western blot analysis. But if this established method is failing us, how should we respond?[58]

As scientific publishers look more closely at experimental methods and data analysis,[18] we have a unique opportunity to re-evaluate our commonly-used normalization methods. These discussions may be uncomfortable, but we cannot simply ignore the issue. A deeper understanding of how our methods work, why they sometimes fail, and how they affect reproducibility is important for choosing the "best" normalization strategy – one that fits the context and biology of the experiment.[58]
Multi-protein normalization strategies such as total protein staining are now emerging as the new standard for immunoblot normalization.[5,20-23,54-57] Membrane staining of sample proteins may be an affordable, reliable

option that easily integrates into your existing laboratory workflow.

Several recent studies hint at the future of Western blot normalization and analysis. Some researchers are making existing internal loading controls more powerful by adding bioinformatics to the mix. Examples include the aggregate analysis strategies recommended by Janes[1] and Degasperi et al.[10] (as discussed in section 3.2). Andrews and Rutherford recently described an online tool for maximum likelihood calibration of immunoblotting data and other assays that measure samples in batches. This "1-step calibration method" computes calibration results iteratively from all measurements, and reduces the sensitivity of results to experimental noise. The data calibration software they describe is open source and in the public domain.[59]

Looking at normalization from new perspectives may help us more fully understand the limitations, and potential strengths, of our existing methods. The future "gold standard" for immunoblot normalization may combine familiar tools with new analytical paradigms that help us more effectively apply the techniques we already use.

9. References

1. Janes KA (2015) An analysis of critical factors for quantitative immunoblotting. *Sci Signaling* 8(371): rs2.
2. Gomes R, Gilda JE and Gomes AV (2014) The necessity of and strategies for improving confidence in the accuracy of Western blots. *Expert Reviews Proteomics* 11(5): 549-60.
3. Taylor SC and Posch A (2014) The design of a quantitative Western blot experiment. *BioMed Res Intl*. doi: 10.1155/2014/361590.
4. Li X, Bai H, Wang X, Li L, Cao Y, Wei J, Liu Y, Liu L, Gong X, Wu L, Liu S, and Liu G (2011) Identification and validation of rice reference proteins for Western blotting. *J Exper Biol.* doi: 10.1093/jxb/err084.
5. Eaton SL, Roche SL, Llavero Hurtado M, Oldknow KJ, Farquharson C, Gillingwater TH, and Wishart TM (2013) Total protein analysis as a reliable loading control for quantitative fluorescent Western blotting. *PLoS ONE* 8(8): e72457.
6. Rocha-Martins M, Njaine B, Silveira MS (2012) Avoiding pitfalls of internal controls: validation of reference genes for analysis by qRT-PCR and Western blot throughout rat retinal development. *PLoS ONE* 7(8): e43028.
7. Pérez-Pérez R, López JA, García-Santos E, Camafeita E, Gómez-Serrano M, Ortega-Delgado FJ, et al. (2012) Uncovering suitable reference proteins for expression studies in human adipose tissue with relevance to obesity. *PLoS ONE* 7(1): e30326.
8. Prokopec SD, Watson JD, Pohjanvirta R, and Boutros PC (2014) Identification of reference proteins for Western blot analyses in mouse model systems of 2,3,4,8-tetrachlorodibenzo-P-dioxin toxicity. *PLoS ONE* 9(10): e110730.
9. Li R and Shen Y (2013) An old method facing a new challenge: re-visiting housekeeping proteins as internal reference control for neuroscience research. *Life Sciences* 92: 747-51.
10. Degasperi A, Birtwistle MR, Volinsky N, Rauch J, Kolch W, Kholodenko BN (2014) Evaluating strategies to normalise biological replicates of Western blot data. *PLoS ONE* 9(1): e87293.
11. Aldridge GM, Podrebarac DM, Greenough WT, and Waller IJ (2008) The use of total protein stains as loading controls: an alternative to high-abundance single protein controls in semi-quantitative immunoblotting. *J Neurosci Meth* 172(2): 250-54.
12. Aguilar HN, Zielnik B, Tracey CN, and Mitchell BF (2010) Quantification of rapid myosin regulatory light chain phosphorylation using high-throughput In-Cell Western assays: comparison to Western immunoblots. *PLoS ONE* 5(4): e9965.
13. Taylor SC, Berkelman T, Yadav G, and Hammond M (2013) A defined methodology for reliable quantification of Western blot data. *Mol Biotechnol* 55: 217-226.
14. Editorial (2004) Gel slicing and dicing: a recipe for disaster. *Nature Cell Biol.* 6(4): 275.
15. Editorial (2013) Enhancing reproducibility. *Nature Methods* 10(5): 367.
16. Gough NR (2015) Focus issue: Tackling reproducibility and accuracy in cell signaling experiments. *Sci Signal*. 8: eg4
17. Bustin SA (2015) The reproducibility of biomedical research: sleepers awake! *Biomol Detec Quant.* 2: 35-42.

18. Announcement (2013) Reducing our irreproducibility. *Nature* 496: 398.
19. Greer S, Honeywell R, Geletu M, Arulanandam R, and Raptis L (2010) Housekeeping genes: expression levels may change with density of cultured cells. *J Immunol Meth.* 355(1-2): 76-9.
20. Moritz CP, Marz SX, Reiss R, Schulenborg T, and Friauf E (2014) Epicocconone staining: a powerful loading control for Western blots. *Proteomics* 14: 162-68.
21. Collins MA, An J, Peller D, and Bowser R (2015) Total protein is an effective loading control for cerebrospinal fluid western blots. *J Neurosci Methods* 251: 72-82.
22. Miller I, Crawford J and Gianazza E (2006) Protein stains for proteomics applications: which, when, why? *Proteomics* 6(20): 5385-408.
23. Gilda JE, Ghosh R, Cheah JX, West TM, Bodine SC, Gomes AV (2015) Western blotting inaccuracies with unverified antibodies: need for a Western Blotting Minimal Reporting Standard (WBMRS). *PLoS ONE* 10(8): e0135392.
24. Mahmood T and Yang P-C (2012) Western blot: technique, theory, and trouble shooting. *North Amer J Med Sciences* 4(9): 429-34.
25. Riis P (2001) *Molecular Biology Problem Solver: A Laboratory Guide*, edited by AS Gerstein. Wiley-Liss, Inc. Chapter 13: 373-97.
26. Schutz-Geschwender A, Zhang Y, Holt T, McDermitt D, and Olive DM (2004) Quantitative, two-color Western blot detection with infrared fluorescence. *LI-COR Biosciences:* 1-7.
27. Gerk PM (2011) Quantitative immunofluorescent blotting of the Multidrug Resistance-associated Protein 2 (MRP2). *J Pharm Toxicol Methods* 63(3): 279-82.
28. Jensen L, Farook MF, and Reiter LT (2013) Proteomic profiling in Drosophila reveals potential Dube3a regulation of the actin cytoskeleton and neuronal homeostasis. *PLoS ONE* 8(4): e61952.
29. Dubitsky A and Perreault S. A model for binding of DNA and proteins to transfer membranes. *Pall Corporation:* Tech Note.
30. Eaton SL, Hurtado ML, Oldknow KJ, Graham LC, Marchant TW, Gillingwater TH, Farquharson C, and Wishart TM (2014) A guide to modern quantitative fluorescent Western blotting with troubleshooting strategies. *J Vis Exper.* 93: e52099.
31. Heidebrecht F, Heidebrecht A, Schulz I, Behrens S-E, and Bader A (2009) Improved semiquantitative Western blot technique with increased quantification range. *J Immunol Meth*. 345: 40-48.
32. Bergendahl V, Glaser BT, and Burgess RR (2003) A fast Western blot procedure improved for quantitative analysis by direct fluorescent labeling of primary antibodies. *J Immunol Meth.* 277: 117-25.
33. Wang YV, Wade M, Wong E, Li YC, Rodewald LW, and Wahl GM (2007) Quantitative analyses reveal the importance of regulated Hdmx degradation for p53 activation. *Proc Natl Acad Sci U S A*. 104(30): 12365-67.
34. Bakkenist CJ, Czambel RK, Hershberger PA, Tawbi H, Beumer JH, Schmitz JC (2015) A quasi-quantitative dual multiplexed immunoblot method to simultaneously analyze ATM and H2AX Phosphorylation in human peripheral blood mononuclear cells. *Oncoscience* 2(5): 542–554.
35. Weldon S, Ambroz K, Schutz-Geschwender A, and Olive DM (2008) Near-infrared fluorescence detection permits accurate imaging of loading controls for Western blot analysis. *Anal Biochem.* 375: 156-58.
36. Baskin, DG and WL Stahl (1993) Fundamentals of quantitative autoradiography by computer densitometry for *in situ* hybridization, with emphasis on 33P. *J Histochem Cytochem* 41(12): 1767-76.
37. Kodak (2007) The essential reference guide for filmmakers: Basic sensitometry and characteristics of film. *Kodak Educational Products*, Code: H-845 CAT No. 145 6144. Eastman Kodak Company.
38. Laskey, R.A (1993) Efficient detection of biomolecules by autoradiography, fluorography or chemiluminescence. Methods of detecting biomolecules by autoradiography, fluorography and chemiluminescence. *Amersham Life Sci. Review* 23: Part II.
39. Gassmann M, Grenacher B, Rohde B, and Vogel J (2009) Quantifying Western blots: pitfalls of densitometry. *Electrophoresis* 30: 1845-55.
40. Butte AJ, Dzau VJ, Glueck SB (2001) Further defining housekeeping, or "maintenance," genes: Focus on "A compendium of gene expression in normal human tissues". *Physiol Genomics*. 21: 7(2): 95-6.
41. Vogel C and Marcotte EM (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genetics.* 13(4): 227-32.
42. Payne SH (2015) The utility of protein and mRNA correlation. *Trends Biochem Sci.* 40(1): 1-3.
43. Tian Q, Stepaniants SB, Mao M, Weng L, Feetham MC, Doyle MJ, Yi EC, Dai H, Thorsson V, Eng J, Goodlett D,

Berger JP, Gunter B, Linseley PS, Stoughton RB, Aebersold R, Collins SJ, Hanlon WA, and Hood LE (2004) Integrated genomic and proteomic analyses of gene expression in mammalian cells. *Mol Cell Proteomics* 3(10): 960-69.

44. Vogel, C, Abreu Rde S, Ko D, Le SY, Shapiro BA, Burns SC, Sandhu D, Boutz DR, Marcotte EM, and Penalva LO, (2010) Sequence signatures and mRNA concentration can explain two thirds of protein abundance variation in a human cell line. *Mol Syst Biol*. 6: 400.
45. Lundberg E, Fagerberg L, Klevebring D, Matic I, Geiger T, Cox J, Algenas A, Lundeberg J, Mann M, and Uhlen M, (2010) Defining the transcriptome and proteome in three functionally different human cell lines. *Mol Syst Biol* 6: 450.
46. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, and Selbach M (2011) Global quantification of mammalian gene expression control. *Nature* 473: 337-42.
47. Wilhelm M, Schlegl J, Hahne H, Moghaddas Gholami A, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S, Marx H, Mathieson T, Lemeer S, Schnatbaum K, Reimer U, Wenschuh H, Mollenhauer M, Slotta-Huspenina J, Boese JH, Bantscheff M, Gerstmair A, Faerber A, and Kuster B (2014) Mass-spectrometry-based draft of the human proteome. *Nature* 509(7502): 582-87.
48. Shapiro-Reznik M, Jilg A, Lerner H, Earnest DJ, Zisapel N (2012) Diurnal rhythms in neurexins transcripts and inhibitory/excitatory synapse scaffold proteins in the biological clock. *PLoS ONE* 7(5): e37894.
49. Bond D, Primrose DA, and Foley E (2008) Quantitative evaluation of signaling events in Drosophila S2 cells. *Biol Proced Online.* 10: 20–28.
50. Georgopoulos NT, Kirkwood LA, Walker DC, Southgate J (2010) Differential regulation of growth-promoting signalling pathways by E-cadherin. *PLoS ONE* 5(10): e13621.
51. Canetta SE, Luca E, Pertot E, Role LW, Talmage DA (2011) Type III Nrg1 back signaling enhances functional TRPV1 along sensory axons contributing to basal and inflammatory thermal pain sensation. *PLoS ONE* 6(9): e25108.
52. Meiler E, Nieto-PelegrÍn E, and Martinez-Quiles N (2012) Cortactin tyrosine phosphorylation promotes its deacetylation and inhibits cell spreading. *PLoS ONE* 7(3): e33662.
53. Gebhart C, Ielmini MV, Reiz B, Price NL, Aas FE, Koomey M, Feldman MF (2012) Characterization of exogenous bacterial oligosaccharyltransferases in Escherichia coli reveals the potential for O-linked protein glycosylation in Vibrio cholerae and Burkholderia thailandensis. *Glycobiology* 22(7): 962-74.
54. Luo S, Wehr NB, and Levine RL (2006) Quantitation of protein on gels and blots by infrared fluorescence of Coomassie blue and Fast Green. *Anal Biochem.* 350(2): 233-38.
55. Welinder C and Ekblad L (2011) Coomassie staining as loading control in Western blot analysis. *J Prot Res.* 10(3): 1416-19.
56. Gilda JE and Gomes AV (2013) Stain-Free total protein staining is a superior loading control to beta-actin for Western blots. *Anal Biochem.* 440: 186-88.
57. Romero-Calvo I, Ocón B, Martinez-Moya P, Suárez MD, Zarzuelo A, Martinez-Augustin O, de Median FS (2010) Reversible Ponceau staining as a loading control alternative to actin in Western blots. *Anal Biochem* 401(2): 318-20.
58. Collins FS and Tabak LA (2014) NIH plans to enhance reproducibility. *Nature* 505: 612-13.
59. Andrews SS and Rutherford S (2015) A novel method and simple on-line tool for maximum likelihood calibration of immunoblots and other measurements that are quantified in batches. *PLoS ONE* submitted; accessed at BioRxiv http://dx.doi.org/10.1101/026005.

LI-COR

4647 Superior Street • Lincoln, NE 68504 • Phone: +1-402-467-0700 • Toll free: 800-645-4267 • Fax: +1-402-467-0819 • biosales@licor.com

Regional Offices

LI-COR GmbH, Germany, *Serving Europe, Africa, and the Middle East.*
LI-COR Biosciences GmbH • Siemensstraße 25A • 61352 Bad Homburg • Germany • Phone: +49 (0) 6172 17 17 771
Fax: +49 (0) 6172 17 17 799 • bio-eu@licor.com

LI-COR Ltd., United Kingdom, *Serving Denmark, Finland, Iceland, Ireland, Norway, Sweden, and UK.*
LI-COR Biosciences UK Ltd. • St. John's Innovation Centre • Cowley Road • Cambridge • CB4 0WS • United Kingdom
Phone: +44 (0) 1223 422104 • Fax: +44 (0) 1223 422105 • bio-eu@licor.com

LI-COR Distributor Network: www.licor.com/bio/distributors • www.licor.com/bio

Doc # 979-15896 12/15