Considerations when quantitating protein abundance by immunoblot

Alicia A. McDonough,1 Luciana C. Veiras,1 Jacqueline N. Minas,2 and Donna Lee Ralph1

1Department of Cell and Neurobiology, Keck School of Medicine, University of Southern California, Los Angeles, California; and 2School of Natural Sciences, University of California, Merced, California

Submitted 16 December 2014; accepted in final form 22 December 2014

McDonough AA, Veiras LC, Minas JN, Ralph DL. Considerations when quantitating protein abundance by immunoblot. Am J Physiol Cell Physiol 308: C426–C433, 2015. First published December 24, 2014; doi:10.1152/ajpcell.00400.2014.—The development of the immunoblot to detect and characterize a protein with an antiserum, even in a crude mixture, was a breakthrough with wide-ranging and unpredictable applications across physiology and medicine. Initially, this technique was viewed as a tool for qualitative, not quantitative, analyses of proteins because of the high number of variables between sample preparation and detection with antibodies. Nonetheless, as the immunoblot method was streamlined and improved, investigators pushed it to quantitate protein abundance in unpurified samples as a function of treatment, genotype, or pathology. This short review, geared at investigators, reviewers, and critical readers, presents a set of issues that are of critical importance for quantitative analysis of protein abundance: 1) Consider whether tissue samples are of equivalent integrity and assess how handling between collection and assay influences the apparent relative abundance. 2) Establish the specificity of the antiserum for the protein of interest by providing clear images, molecular weight markers, positive and negative controls, and vendor details. 3) Provide convincing evidence for linearity of the detection system by assessing signal density as a function of sample loaded. 4) Recognize that loading control proteins are rarely in the same linear range of detection as the protein of interest; consider protein staining of the gel or blot. In summary, with careful attention to sample integrity, antibody specificity, linearity of the detection system, and acceptable loading controls, investigators can implement quantitative immunoblots to convincingly assess protein abundance in their samples.

immunoblot; Western blot; antibody-antigen; immunodetection; quantitation

IN 1979, RENART, REISER, AND STARK (22) reported a method to rapidly detect proteins as antigens after their resolution by native or SDS-PAGE. The procedure involved passive transfer of the proteins out of the gel onto paper coated with diazobis groups that would covalently bind proteins as they were passively wicked out of the acrylamide gel. After blocking the remaining sites, the blot was incubated with diluted antibody, then with radiolabeled protein A, and finally analyzed by autoradiography. This method was a breakthrough for investigators because it not only provided a method to find a “needle in a haystack,” that is, a low-abundance protein in a homogenate, but also provided a method to efficiently screen a large number of antibodies, e.g., hybridoma samples, against a specific antigen. Although cumbersome (required coating paper with diazo groups and iodinating protein A), many labs used this method in the 1980s to address questions about protein subunit expression, assembly, and regulation; our own lab used these blots to study sodium pumps in a variety of tissues (16–18, 24). Around the same time, other labs reported simpler procedures that involved electrophoretically blotting the gel onto nitrocellulose (5, 26), and working in the Pacific Northwest, Burnette coined the term “Western blot,” in homage to Edwin Southern’s development of the DNA blot. Subsequent immunoblot simplifications included the use of secondary antibodies tagged with chemiluminescent or fluorescent probes to eliminate the need for radioactivity. The immunoblot is now, 35 years later, a ubiquitous tool used in most all aspects of biological research as well as a tool for commercial immunodiagnostics (4).

While the Western blot was not originally designed to quantitate protein abundance with antibodies, this approach is quite commonly implemented to do just that. Our experience as investigators, reviewers, and readers has revealed a wide disparity in the rigor of execution and interpretation of immunoblot quantitation by authors and a similarly disparate rigor of reviewers’ attention to critical aspects of quantitation. The aim of this short review is to consider several issues that we have found to be important for immunoblot quantitation, including considering sample integrity and characteristics, demonstrating antibody specificity, establishing the linearity of the detection system, and implementing acceptable loading controls.

Our report will not discuss how to optimize immunoblot signal-to-noise ratio, nor will we rank one detection and quantitation method versus another. There are many excellent resources available that describe how to perform and optimize immunoblots, including excellent vendor documents available online, reviews, and books (2, 12, 14). Our goal is to provide

Address for reprint requests and other correspondence: A. A. McDonough, Cell and Neurobiology, Keck School of Medicine of USC, 1333 San Pablo St. MMR 508, Los Angeles, CA 90033 (email: mcdonoug@usc.edu).
an introduction to important considerations in quantitating immunoblot signals for both novices and experts as well as for reviewers and critical readers. We will use examples from tissues, but the same considerations apply to cell culture samples.

Consider Sample Integrity and Properties

Is the analysis comparing equivalent samples? Protein abundance is routinely normalized to the amount of protein loaded per lane (more on determining the actual amount to load is discussed below). If the experimental aim is to determine how a protein of interest changes with a treatment, disease process, genotype, or aging, it is important to consider confounding factors. Figure 1A provides images of normal and infarcted myocardium samples (at approximately the same magnification) to illustrate this point. About 50% of the cardiac muscle tissue is replaced by fibrotic tissue in the infarcted sample. Said another way, the total protein that will be used to normalize the relative abundance of the protein of interest is now made up of about 50/50 muscle and fibrotic tissue. Consequently, even if there is not a change in expression of the protein of interest in the muscle cells in the infarcted sample, its relative abundance will decrease about 50% when compared to normal myocardium. When studying highly fibrotic or otherwise modified tissues, the investigator must carefully consider the question under investigation and determine the best way to normalize the relative abundance of the protein of interest, for example, the abundance of a muscle-specific structural protein could be assessed in parallel by immunoblot and used to normalize the protein of interest. On the other hand, if the goal is to quantitate the actual degree of fibrosis, then normalizing a fibrotic marker to a constant amount of sample protein could be appropriate.

Obtaining relative abundance data usually requires accumulating a sufficient number of samples per group and, ideally, analyzing the groups on a single blot (with multiple gels on one blot paper if necessary) so that all the samples to be compared are incubated with the same aliquots of primary and secondary antibodies, and quantified simultaneously. These considerations reduce the variables that interfere with quantifying relative abundance in the groups. If all the samples cannot be collected and prepared on the same day, then the sample handling between collection and assay becomes an important variable that warrants serious consideration. Often tissues are quick frozen and then prepared for assay together. Figure 1B illustrates the results of a comparison of samples of mouse kidney homogenates prepared from kidneys quick frozen in liquid nitrogen and stored for a month at −80°C, versus kidney homogenates prepared immediately after euthanizing the mice. Results are shown for the renal Na⁺/H⁺ exchanger isoform 3 (NHE3, the sodium transporter responsible for the bulk of the renal sodium reabsorption). The apparent abundance of NHE3 is less in the frozen versus the freshly prepared homogenates, indicating that there has been a partial loss of the epitope (perhaps by degradation or masking) as a consequence of freezing and thawing. The mouse samples, in our opinion, are still useful as long as they are prepared identically (all freshly prepared or all frozen for about the same time). However, our experience indicates that frozen rat kidneys do not fare as well as frozen mouse kidneys. Figure 1C illustrates samples from frozen and fresh rat kidneys prepared with the same reagents.
on the same day and analyzed on the same blot. While there are strong signals for the NHE3, the Na\(^+\)-K\(^+\)-2 Cl\(^-\) cotransporter (NKCC), and Na\(^+\)-Cl\(^-\) cotransporter (NCC) in the freshly prepared homogenates, the signals are lost in kidneys that were freeze/thawed before homogenate preparation. For this reason, we recommend that the investigator always compare homogenates made from freshly isolated tissue to frozen tissue to understand if there is serious loss of antibody epitope for their protein of interest. Interestingly, we have found that once renal homogenates are prepared with protease and phosphatase inhibitors (19) and stored in single-use aliquots at \(-80^\circ C\), the immunoblot signals of renal transporters do not appear to further decay over many years of storage, whether the samples were originally from fresh or freeze/thawed tissue (not shown).

Many sample handling factors can affect relative abundance. As two examples, the phosphorylation status of the renal NCC is profoundly influenced by the time since consuming the last meal (25), and we have now been made aware of the influence of circadian rhythms on protein expression (23). In addition, sample enrichment by subcellular fractionation or affinity purification will exhibit batch-to-batch variability in relative recoveries; analysis of homogenates circumvents this recovery consideration.

In summary, the investigator needs to consider the status of the tissues being compared and assess how handling between collection and assay can influence the calculation of a sample or treatment groups’ relative abundance.

**Consider the Specificity of the Antibody for the Target**

It is the investigator’s responsibility to provide details about each antiserum sufficient for a reader to replicate the immunodetection. This requires providing not just the vendor information but also catalog number, because the vendor can have multiple antibodies to one protein. Information about antiserum dilution, host, and specific secondaries are also very useful because antibody vendors come and go while an antisera can be used by an author over decades. If many antisera are used, these can be organized in a table.

Does the antiserum recognize the protein of interest? It is the investigator’s responsibility to answer this question. It goes without saying that every immunoblot needs to include molecular weight standards to assess and report the apparent molecular weight of the protein(s) detected by the antiserum. It is important to acknowledge that just because a vendor advertises that an antiserum recognizes a target protein does not guarantee that the antibody detects that protein in your samples, nor that it does not also bind to another unrelated protein (nonspecifically) with a similar molecular weight. As a case in point, a recent study by Herrera, and a companion commentary by the Eguchi lab, provide proof that seven commercially available antibodies to the angiotensin type 1 receptors (AT1R) lack specificity (7, 11). Bands of the “expected” apparent molecular weight were equally evident in tissues and cells that expressed no AT1R as in tissues known to express AT1R. The invest-

---

**Fig. 2. Consider the specificity of the antibody for the target.** SPAK [sterile 20 (STE20)/SPS1-related proline/alanine-rich kinase] was detected first with an anti-NH\(_2\)-terminal epitope antiserum, which detects full-length SPAK (FL-SPAK), then a red-tagged secondary; the blot was reprobed with an anti-COOH-terminal epitope antiserum that detects FL-SPAK, smaller SPAK 2, and an even smaller kidney-specific SPAK isoforms (KS-SPAK), then a green-tagged secondary. **Bottom:** illustration of the image obtained when both signals were detected at the same time; yellow indicates detection at the same location. SPAK is expressed in testes as a single full-length isoform (FL-SPAK) included as a positive control, is absent in a SPAK knockout mouse (SPAK-KO), included as a negative control, and all three isoforms are present in wild-type mouse kidney. Nonspecific background bands (white arrows) were identified based on expression in the SPAK-KO. See text for details and discussion.
gators postulate that this may be a common problem with G protein-coupled receptors and provide the opinion that authors (and vendors) must apply rigorous standards to establish the specificity of the antibody for the target. Given the significant effort required to characterize an antibody, as well as potential pitfalls, unbiased resources have begun to be established on the web to aid investigators in choosing antibodies that have a good chance of working based on previous published reports (e.g., www.Citeab.com) and epitope sequence considerations (e.g., www.uniprot.org).

We provide a typical characterization for the transporter-activating kinase known as SPAK (sterile 20 (STE20)/SPS1-related proline/alanine-rich kinase), using antibodies to COOH- and NH2-terminal epitopes, in Fig. 2. These antibodies, prepared by the Delpire lab for use in studying brain SPAK (20, 21), were applied by us to the study of kidney SPAK. Unlike brain SPAK, in the kidney, SPAK is expressed as three nested isoforms sharing a common COOH terminus: full-length SPAK (FL-SPAK), SPAK 2, and a smaller kidney-specific form (KS-SPAK) (15). The Delpire lab generously provided our lab with both antibodies as well as a kidney from a SPAK knockout (KO) mouse, to serve as a negative control. Initial characterization of the antibodies (20) revealed very strong SPAK signals in testis, which serves as a positive control. The top panel (Fig. 2) displays the immunoblot results obtained after probing with the anti-NH2-terminal epitope antibody, present only on FL-SPAK, followed by a secondary antibody with a red label: The testis sample lane gives a strong positive FL-SPAK signal, there is no staining in the SPAK KO lane in the FL-SPAK region, and strong bands are evident at the position of FL-SPAK in the two wild-type (WT) mouse kidney sample lanes. However, there are three prominent bands below the FL-SPAK band coincident in both the WT and KO samples, indicating that these are nonspecific “background” bands. The middle panel displays the results obtained when the same blot was probed with an anti-COOH-terminal epitope antiserum expected to detect all three of the nested renal SPAK isoforms, followed by a secondary antibody with a green label: A single strong band is evident in the positive control lane containing testis (which expresses only FL-SPAK), only a very faint background band is evident in the SPAK KO lane, indicating that this antiserum is quite specific, and multiple bands are evident in the two WT lanes. The two images were superimposed, which indicates that 1) the FL-SPAK, recognized by both the anti-COOH- and anti-NH2-terminal antisera, appears yellow (indicating that they both recognize the same protein) in testis and the WT samples; 2) the SPAK 2 and KS-SPAK, which are recognized by only the COOH-terminal probe, appear green; and 3) “background bands,” detected by the NH2-terminal antibody, appear red. This example illustrates the utility of using fluorescent tagged secondary antibodies to multiple epitopes to discriminate between specific and non-specific labeling. With this information we were able to identify SPAK, SPAK 2 and KS-SPAK and determine how their levels were regulated by angiotensin II as we observed for renal NCC using nonnondenaturing gels (13). When using anti-phospho-specific antiserum, the main concern is that the antiserum may also cross react with the nonphosphorylated protein. To avoid this, it is recommended to preabsorb the antiserum with a nonphosphorylated version of the immunizing peptide and to keep the primary antiserum incubation time short to minimize cross-reaction with the nonphosphorylated protein.

In summary, the investigator needs to convincingly establish the specificity of the antiserum for their protein of interest by assessing whether it is detected at the correct apparent molecular size (by providing a stained lane of molecular weight markers on the immunoblot); by determining whether it is present in a known positive control (often available for purchase from vendors); by determining whether it is absent in a sample that does not contain the protein of interest; and by identifying nonspecific “background” bands.

Consider the Linearity of the Detection System

How much protein should be assayed? Assuming that the integrity of the sample and specificity of the antiserum for the

![Fig. 3. Consider the linearity of the detection system.](image)

Fig. 3. Consider the linearity of the detection system. A: illustration of the immunodetection of sodium pump catalytic isoform subunits (α1, α2, α3) and the Na+/Ca2+ exchanger (NCE) from human heart detected between 20 and 100 µg/lane. [Redrawn from Wang et al. (27).] B: illustration of the immunodetection of sodium pump (NKA) α1-subunit between 0.12 and 0.5 µg/lane and β1-subunit between 3 and 48 µg/lane in rat kidney homogenates from cortex and medulla. See text for details and discussion.
protein of interest are satisfactorily characterized, the investigator and critical reader should consider whether the conditions of sample loading and quantitation provide signals in the linear range of the detection system (regardless of whether the system is detecting light or radioactivity or fluorescence). That is, one should determine whether the signal generated from the primary and secondary antibody labeling is a direct measure of the relative abundance of the protein of interest in the samples. It is our opinion that linearity of the detection system should be established for each protein of interest in every study (ideally in every immunoblot) because of the high number of variables that can influence the final signal. Variables include the level of protein expression in the tissue, sample handling and preparation (Fig. 1), gel transfer conditions, blocking efficiency, antibody dilution, affinity for the target, and detection method (e.g., chemiluminescence emits a signal that changes with time); we contend that significant regulation is often missed or inaccurately assessed because immunoblots are analyzed without assessing linearity. We propose that investigators analyze multiple amounts of samples on the same blot to calibrate detection; Charette et al. (6) propose a similar strategy emanating from their analysis of nonlinear chemiluminescent signals. The inset of Fig. 3A provides a series of immunoblots of cardiac homogenates in which samples were loaded at 20, 40, 60, 80, and 100 µg/lane then individually probed with antisera directed against sodium pump catalytic subunit isoforms (α1, α2, and α3) or against the sodium/calcium exchanger (NCE) (redrawn from Ref. 27). The graph (Fig. 3A) plots relative abundance of the immunoblot signal from each protein of interest, normalized to the abundance of the signal from the 60 µg sample defined as 1.0; the dotted line indicates equivalent changes in protein loaded and density signal. Between 20 and 80 µg/lane, the change in relative signal abundance is roughly equivalent to the amount of protein loaded for α2, α3, and NCE, but not for α1. This is a good illustration of a problemmatic issue because while the density of the α1 signal is “light,” compared with the NCE signal, the detection system is saturated above 60 µg for α1, but not for NCE. Said another way, just because signals on an immunoblot are in the gray zone does not guarantee they are in a linear range of detection—a doubling of signal with abundance must be validated for the specific assay conditions.

Figure 3B compares detection of renal sodium pump subunits (α1 and β1). Despite the fact that the analysis shown is restricted to a linear range, it is clear that this linear range of detection for kidney α1 is attained with about 100-fold less homogenate than found for cardiac tissue, corresponding to the far greater abundance of Na-K-ATPase α1 in kidney versus heart. While this finding is not surprising, Fig. 3B also indicates that detection of the Na-K-ATPase β1 subunit in renal tissue is linear when loaded at 100-fold more than that for detection of α1. Why are the ranges not more similar for this αβ heterodimer? One explanation is that β1 is a glycoprotein and the sugar residues reduce antibody-antigen binding; we and others established long ago that β detection improves when β is deglycosylated before immunodetection (3). Another explanation that needs to be considered for all immunodetection is that an antisera’s affinity for an antigen-containing protein is dependent on the biological reaction of the immunized host against the injected antigen during antisera production and, thus, is highly variable between hosts or hybridoma clones.

Fig. 4. Consider the loading controls: part 1. A: illustration of a single immunoblot constructed from two SDS-PAGE gels loaded with 6 µg/lane (top of blot) and 3 µg/lane (bottom part of blot) kidney homogenate and probed with antiserum to claudin 10 and a secondary antiserum tagged red. The quantitation, indicated to the right of the blot, demonstrates linearity of the detection system: Specifically, signal intensity increases 1.7-fold with sample doubling. B: Image of the same blot reprobed with an anti-actin (pan-actin) antiserum and a secondary antiserum tagged green. The actin signal intensity does not change with sample doubling. We conclude that actin is not an acceptable loading control for this application; rather it is “overloaded” at 3 and 6 µg kidney homogenate/lane. See text for further details and discussion.
In summary, if an investigator aims to quantify changes in protein expression (or phosphorylation) using immunoblots, it is our experience and opinion that each immunoblot should contain a demonstration of the linearity of the detection system. This can be accomplished in a number of ways. We routinely run samples at 1 and ½ protein amounts (which may require multiple gels that are transferred together on the same blot membrane) and then process the single blot through immunodetection using the identical primary and secondary antibodies and simultaneous quantitation (as in Fig. 4). We calculate the linearity ratio by dividing the signal obtained with the full amount with signal obtained with ½ the amount and expect it to be close to 2. If not, then the assay is underestimating or overestimating the effect of the treatment. In some cases we have included blots containing both 1 and ½ loading amounts in our figures (19) and in others we have reported the method used to establish linearity and shown just one amount in the figures (9). In this latter case, we have been asked by reviewers to provide images of every immunoblot loaded at the two amounts for verification. Another option is to provide a linearity analysis (as in Fig. 3A) as a supplemental figure, if the journal allows.

**Consider the Loading Controls**

Is a loading control necessary if one has established the linearity of the detection system? Yes. Equivalent loading of sample amount per lane is important to establish if immunoblots will be quantitated. Unequal sample-to-sample loading can result from: high levels of variability in the determination of protein concentration of the samples, incorrect sample preparation, or poor pipetting into the wells. It is very common for investigators to reprobe their blots with what is known as a loading control or housekeeping protein (e.g., actin or GAPDH) that exhibits high-level, constitutive expression in the tissue and is not expected to vary with the treatment. However, a key problem with this approach is that these loading control proteins are rarely in the same linear range of detection as the target protein of interest. A number of studies concur that in the typical loading range of 10–50 µg of cell lysate, quantitation of housekeeping proteins by immunodetection is not possible because they are in saturating quantities (1, 10, 28). A good illustration of the issue is provided in Fig. 4 in which five samples of renal homogenates were resolved at 6 and 3 µg/lane (to establish linearity). The immunoblot was first probed with an antiserum directed to claudin 10 (a tight junction protein) and a red-tagged secondary (Fig. 4A, left) and claudin 10 was detected between the 17 and 28 kDa molecular mass markers. Quantitation of the five samples, provided to the right, indicates a good linearity of the detection system as the signals were roughly doubled with 6 versus 3 µg protein loading. The same blot was probed with an antiserum to actin and a green-tagged secondary, and actin was detected above 34 kDa (Fig. 4B). The signal intensity of the actin band does not significantly change when sample loaded was reduced to ½; thus, actin is not an acceptable loading control for this application; it is saturating the membrane at just 3 µg/lane.

An alternative to a single loading control is protein staining of the gel or blot as a loading control (1, 28). Vendors are starting to market kits to stain loading gels, but it can be done with what is on hand in the lab. We make up enough of each sample to run a parallel “loading gel” which is loaded and run identically to the gel(s) that will be immunoblotted (another approach is to actually stain the blot but we are concerned that this could interfere with immunodetection). In our approach, used by others as well, the gel is not blotted, rather, it is simply stained and destained and a number of bands are chosen to quantitate by densitometry. Another approach is to quantitate the density of the entire lane from top to bottom, but we have had good results assessing discrete bands that do not appear to be too dark or too light. Figure 5 provides an illustration of a loading gel containing eight samples that were all, nominally, loaded with 7.5 µg/lane of renal cortex homogenate. Five bands were identified, quantitated by densitometry, then normalized to the mean density for that band in all eight samples, defined as 1.0. The relative abundance results for each sample as well as the average value for the five bands within each sample are displayed below the loading gel (Fig. 5). A visual inspection of the loading gel suggests samples 4 and 7 are relatively overloaded and the quantitation bears this out, as both have relative average densities of 1.13; the others vary by 5% or less. At this point, the investigator can report the level of variability that is acceptable for the quantitative immuno-
blots, and options for proceeding include correcting the density of the immunoblots for this variance, beginning again with the protein assays and sample preparation to reduce sample to sample variability to 5–10% from the mean average density or less, or reporting and accepting the variability after providing a rationale. In summary, staining the gel provides a direct image of the actual protein loaded. The stained gel approach is a far more valuable loading control than that obtained by quantitating saturated signals from housekeeping proteins. The latter reveals little more than whether the sample was loaded or not.

Conclusion

During the past 35 years, immunoblotting has been a critical tool for making discoveries and progress in many areas of investigation from early use in expression cloning of genes, to utilization for studies of protein expression, to screening antibodies and clinical immunodiagnostics. Physiologists, with their focus on how tissues and organ systems work normally and in response to challenges, have relied on the Western blot’s ability to detect a “needle in a haystack” to address questions about organ system operation and regulation. However, since the method is simple to set up and implement, results are often conducted and reported without the requisite controls. Samples and reagents must be validated and controlled to obtain interpretable results. Said another way, if an investigator aims to determine whether there are more needles in one haystack than another, it is key to convincingly validate the integrity of the sample, the specificity of the antibody, and the linearity of the detection system and to assess sample loading.

NOTE ADDED IN PROOF

We recently became aware of a very useful review by Murphy and Lamb (18a) that similarly reviews important considerations for quantitating proteins in muscle using antibody-based techniques.

ACKNOWLEDGMENTS

Drs. Jo Adams and Christina Bennett provided very helpful comments.

GRANTS

This work was supported by National Institute of Diabetes and Digestive and Kidney Diseases Grant DK-083785 (to A. A. McDonough).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS


REFERENCES


