"This is the peer reviewed version of the following article: Carbonara, K, Padula, MP, Coorssen, JR. Quantitative assessment confirms deep proteome analysis by integrative top–down proteomics. *Electrophoresis*. 2023; 44: 472– 480, which has been published in final form at 10.1002/elps.202200257. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited."

Quantitative assessment confirms deep proteome analysis by integrative top-down proteomics

Katrina Carbonara,[1] Matthew P. Padula,[2] and Jens R. Coorssen,[3,4*]

[1]Department of Health Sciences, Faculty of Applied Health Sciences, Brock University, St. Catharines, Ontario, Canada; kcarbonara@brocku.ca.

[2] School of Life Sciences and Proteomics, Lipidomics and Metabolomics Core Facility, Faculty of Science, University of Technology Sydney, NSW, Australia.

[3]Department of Biological Sciences, Faculty of Mathematics and Science, Brock University, St. Catharines, Ontario, Canada; jrcoorssen@gmail.com.

[4]Ronin Institute, Montclair, NJ, USA.

*Author to whom correspondence should be addressed

Keywords:

- Proteoforms
- Isoelectric focusing
- IPG strips
- Two-dimensional gel electrophoresis

Abbreviations:

1DB – 1DE sample buffer

1DE – 1-dimensional gel electrophoresis (SDS-PAGE)

2DB – 2DE sample buffer

2DE – 2-dimensional gel electrophoresis

AU – arbitrary units

LC/MS/MS – liquid chromatography and tandem mass spectrometry

nIRFD – near-infrared fluorescence detection

Abstract

The goal of integrative top-down proteomics (i.e., two-dimensional gel electrophoresis (2DE) coupled with liquid chromatography and tandem mass spectrometry (LC/MS/MS)) is a routine analytical approach that fully addresses the breadth and depth of proteomes. To accomplish this, there should be no addition, removal, or modification to any constituent proteoforms. To address two-decade old claims of protein losses during front-end proteome resolution using 2DE, here we tested an alternate rehydration method for immobilized pH gradient strips prior to isoelectric focusing (IEF; i.e., face-up compared to face-down) and quantitatively assessed losses during the front-end of 2DE (rehydration and IEF). Using a well-established high-resolution, quantitative 2DE protocol, there were no detectable proteoform losses using the alternate face-up rehydration method. While there is a <0.25% total loss of proteoforms during standard face-down rehydration, it is insignificant in terms of having any effect on overall proteome resolution (i.e., total spot count and total spot signal). This report is another milestone in integrative top-down proteomics, disproving long-held dogma in the field and confirming that quantitative front-end 2DE/LC/MS/MS is currently the only method to broadly and deeply analyze proteomes by resolving their constituent proteoforms.

1 Introduction

Rather than simply inferring the presence of canonical amino acid sequences, the goal of top-down proteomics is to provide the broadest and deepest possible quantitative analysis of a given proteome [1-5]. To achieve this, there should be no loss of proteoforms between sampling and analysis. Currently, the gold standard for resolution, detection, and quantification of proteoforms is two-dimensional gel electrophoresis (2DE; resolving proteoforms first by charge (isoelectric focusing; IEF) then by nominal molecular weight (SDS-PAGE)) coupled with liquid chromatography and tandem mass spectrometry (LC/MS/MS). This integrative top-down approach is currently the only method enabling the necessary breadth and depth of proteome analyses (i.e., fully resolving and identifying proteoforms). Only such rigorous quantitative analyses can provide genuine insight into molecular mechanisms and identify critical biomarkers and therapeutic targets [1, 3, 4]. Nonetheless there is no

‘perfect’ method, including 2DE. As such, our ongoing goal has been further refinement of 2DE by critically assessing and thus improving the different stages of the analytical process [6-16].

Since the pioneering innovation of immobilized pH gradient (IPG) strips to better enable the first dimension step (IEF) of 2DE [17], there has been limited assessment and refinement of this critical process that involves rehydration/loading of the strips with a buffer + protein sample/proteome extract prior to IEF. In-gel rehydration, in which sample is mixed into the rehydration buffer and uniformly introduced to the entire IPG strip in the bottom of a tray with narrow wells (i.e., strip is placed face-down into the sample), is by far the most efficient and widely used approach [18]. While this and other refinements improve resolution over the entire pH and molecular weight (MW) range [7], there have been reports of proteins adhering to the walls of the rehydration chamber during overnight rehydration, resulting in apparent marked losses of the total protein applied [19, 20]. Additionally, it was also reported that a notable fraction was additionally lost to the paper wicks and mineral oil used, respectively, to effect electrical contact and sample desalting, and to prevent oxidation of proteoforms during IEF [20].

The objective here was to further assess those apparent losses during IEF and identify an alternative rehydration method that would reduce potential proteoform losses. We propose a face-up rehydration method during which the rehydration buffer and proteome sample do not come in contact with the rehydration tray. Additionally, we again evaluated losses to wicks and mineral oil during IEF following both standard (face-down) and face-up rehydration of IPG strips.

2 Materials and Methods

All chemicals were of ultra-pure or analytical grade. ReadyStrip™ IPG strips (7 cm, 3-10 NL), Bio-Lyte carrier ampholytes (pH 3-10, 3-5, 4-6, 6-8, 7-9), and unstained broad range (10-250 kDa) protein standards were from Bio-Rad Laboratories (Hercules, CA). Sodium chloride, methanol, and mineral oil (light) were from Fisher Scientific (Hampton, NH). Acrylamide and citric acid were from BioShop Canada Inc. (Burlington, ON). Sodium orthovanadate, sodium fluoride, pepstatin A, benzamidine hydrochloride hydrate, thiourea, urea, N,N’-methyl-enebisacrylamide, acrylamide/bisacrylamide solution (40%, 37.5:1), N,N,N’,N’-

tetramethylethylenediamine, sodium n-dodecyl sulfate, tri-n-butylphosphine, ammonium sulfate, glycerol, phosphoric acid, ammonium peroxydisulfate, and ammonium hydrogen carbonate were from Alfa Aesar (Ward Hill, MA). PBS tablets were from Medicago (Uppsala, SE). Coomassie brilliant blue G-250, CHAPS, AEBSF, leupeptin, agarose I, TG-SDS buffer (powder), tris, and tris hydrochloride were from VWR (Radnor, PA). Aprotinin, staurosporine, tricine, acetonitrile, and formic acid were from Sigma-Aldrich (St. Louis, MO). Acetic acid and hydrochloric acid were from Anachemia (Mississauga, ON). PageRuler™ unstained low range protein ladder and DTT were from Thermo Fisher Scientific (Waltham, MA). Samples of previously snap-frozen whole rat brain (*Rattus norvegicus*) were homogenized via automated frozen disruption as previously described [9, 15]. The resulting powdered total rat brain was solubilized in 2DE sample buffer (2DB; 8 M urea, 2 M thiourea, 4% (w/v) CHAPS, and 1x protease inhibitors supplemented with kinase/phosphatase inhibitors) [7]. Aliquots were snap-frozen and stored at -80°C until used. 2DE was carried out as previously described [15, 16]. Total protein concentration was assessed using a solid-phase assay as described [21].

2.1 2DE

Reduction and alkylation were carried out at 25°C (Digital Dry Bath (Bio-Rad, Hercules, CA)) with 45 mM/2.3 mM DTT/TBP for one hour, followed by 230 mM acrylamide for one hour [7, 22] (Butt 2005, Butt 2007) immediately prior to overnight (i.e., 16 h) IPG strip (7cm 3-10NL (Bio-Rad, Hercules, CA)) rehydration, similar to the studies of interest [19, 20]. Rehydration was done either gel face-down or gel face-up with either 2DB alone or 2DB with 100 μg/125 μL of total rat brain protein. This was done in a i12™ rehydration/equilibration tray (Bio-Rad, Hercules, CA). Face-up rehydration was done by carefully beading the sample along the entire length of the strip. Following the 16 h rehydration period, IPG strips were removed and 1 mL of 2x1DE buffer (2x1DB; 50 mM Tris [pH 8.8], 4% (w/v) SDS, 100 mM DTT, 15% (w/v) glycerol, 0.002% (w/v) bromophenol blue, and 1x protease inhibitors) was incubated in each of the wells of the rehydration tray (i.e., wells that had been empty as well as parallel wells that had contained IPG strips) for 30 min at RT with occasional agitation. The resulting 2x1DB samples were aliquoted and stored at -80°C.

During the first 2.5 h of IEF (PROTEAN® i12™ IEF System (Bio-Rad, Hercules, CA)), wicks (Whatman 3MM cellulose chromatography paper (GE Healthcare, Buckinghamshire, UK)) were routinely changed 5 times (i.e., 15 min, 45 min, 1.25 h, 1.75 h, 2.25 h). Wicks from the anodic and cathodic ends were removed and stored in a 50 mL conical tube. Unused wicks from the same batch of paper were stored in a separate 50 mL conical tube. Wicks were then incubated with ~1 mL of 2x1DB per wick. Samples were vortexed at 0 min, 15 min, 30 min, and every 30 min for 2.5 h to ensure recovery of any proteoforms that might be present. The resulting solution was recovered, aliquoted, and stored at -80°C until analysis.

Following IEF, mineral oil was removed from the tray and combined with equal volumes of 2x1DB. Additionally, unused mineral oil was also combined with equal volumes of 2x1DB. Samples were vortexed at 0 min, 15 min, 30 min, and every 30 min for 2.5 h to ensure recovery of any proteoforms that might be present. The resulting solution was recovered, aliquoted, and stored at -80°C until analysis. IPG strips were incubated in equilibration buffer (6 M urea, 0.375 M tris (pH 8.8), 2% (w/v) SDS, 10% (w/v) glycerol), supplemented with 2% (w/v) DTT for 10 min, followed by 350 mM acrylamide for 10 min. IPG strips were then overlayed onto 5% T stacking gel and embedded in 0.5% (w/v) low-melting agarose prior to SDS-PAGE and carried out as described below (**2.2**).

2.2 Tris Glycine SDS-PAGE

SDS-PAGE was carried out as previously described [15] with minor modifications. Resolving gels (1.0 mm thick 12.5% T; Bio-Rad mini-gel format) were overlaid with a 5% T stacking gel with 5 mm wide wells. Samples in 2x1DB from the rehydration tray, wicks, and mineral oil were diluted to 1DB, vortexed, heated at 100°C for 5 min, bath sonicated for 5 min, and cooled to RT prior to loading into the stacking gel wells. All samples were resolved in parallel and included: 2 μg/5 μL of total rat brain protein sample, 5 μL of extracts (mineral oil, wicks, rehydration tray), and 2 μL of unstained broad range (10-250 kDa) protein standards (Bio-Rad). Electrophoresis was carried out at 4°C at 150 V for 12 min (i.e., until samples had fully entered the stacking gel) and then 90 V to completion (PowerPac Universal Power Supply (Bio-Rad, Hercules, CA)). Following resolution, gels were fixed with 1 M citric acid + 5% (v/v) acetic acid [16] for 1 h at RT with gentle rocking followed by 3x20 min washes

with ddH_2O with gentle rocking. An established high sensitivity colloidal Coomassie Brilliant Blue (cCBB) staining protocol was carried out for 20 h as previously described [11, 15]. Gels were subsequently destained 5x15 min with 0.5 M NaCl and imaged using near-infrared fluorescence detection (nIRFD; 685/≥750 nm ex/em) on an Amersham Typhoon 5 Biomolecular Imager (Cytiva Life Sciences, Marlborough, MA) [10-12]; pixel size was set to 50 μm and PMT gain to 600 V.

2.3 Tris-tricine SDS-PAGE

A modified Tris-tricine gel system was used to resolve lower molecular weight species, as previously described [7, 23]. SDS-PAGE was carried out using 1.0 mm thick 15-20% T gradient resolving gels (BioRad mini-gel format) overlaid with a 5% T stacking gel with 5mm wide wells. Samples from the rehydration tray, solubilized wicks, and mineral oil in 1DB were vortexed, heated at 100°C for 5 min, sonicated for 5 min, and cooled to RT prior to loading into the stacking gel wells. All samples were resolved in parallel and included: 2 μg/5 μL of total rat brain protein sample, 5 μL of extracts (mineral oil, wicks, rehydration tray), and 2 μL of unstained low range (3.4-100 kDa) protein standards (Thermo Fisher Scientific, Waltham, MA). Electrophoresis, fixing, staining, and imaging were carried as described in section 2.2.

2.4 Preparative gels

As a direct quantitative assessment of the original total protein load used in the 2DE analyses, SDS-PAGE was carried out as previously described [15] with minor modifications. Resolving gels (1.0 mm thick 12.5% T; Bio-Rad mini-gel format) were overlaid with a 5% T stacking gel having a ~7 cm comb (i.e., to mimic an IPG strip). 100 μg of total rat brain protein originally solubilized in 2DB was supplemented with 1DE buffer (1DB; 25 mM Tris [pH 8.8], 2% (w/v) SDS, 50 mM DTT, 7.5% (w/v) glycerol, 0.001% (w/v) bromophenol blue, and 1x protease inhibitors) and loaded onto the wide well of the stacking gel. Electrophoresis, fixing, staining, and imaging were carried out as described in section 2.2.

2.5 Data analysis

Intensity of lane signals on 1D gel images (16-bit tiff) were analyzed using ImageLab (v. 6.1.0; Bio-Rad Laboratories Inc.). Using the lane volume tool, a representative lane was outlined, and this outline was then copied to remaining lanes

of the gel. Two background measurements were taken from immediately adjacent, protein-free areas of the gel. Data were analyzed using Excel (v. 16.58) and Prism (v. 9.3.1; GraphPad Software); $p < 0.05$ was considered significant. Signal intensity was quantified as volume (the sum of the pixel volumes within the specified areas) in arbitrary units (AU; ImageLab v. 6.1.0) and normalized to yield signal/mm^2. Average local background was subtracted from corresponding lane signals. Errors are reported as standard error of the mean (SEM).

To assess 100% load of rat brain protein in parallel with other 1DB samples, it was impractical to load 100 μg of total protein in a single lane; thus, 2 μg of total protein was loaded. The resulting signal intensity was multiplied by 50 to obtain a theoretical 100 μg load (100%). All other lane signal intensities from all sample types were multiplied by 50 to normalize these to the 100% load of rat brain total protein.

All 2D gels and preparative gels were analyzed using Delta2D (v. 4.8.2; DECODON), as previously described [15, 16]. Images were warped and automated spot detection was carried out; artefacts were manually excluded from analysis using the spot editing tool. The final spot 'pattern' was applied to all 2D gel images to obtain absolute spot volumes. Analysis criteria were as follows: false discovery rate (FDR) = 1%, relative standard deviation ≤30% of spot signals between replicates, and fold changes of ≥2 (increase or decrease) relative to the control ($p < 0.05$ was considered significant). Individual gel spot count was determined using automated spot detection on raw, unfused images.

2.6 In-gel extraction

Bands of interest were manually excised from the triplicate gels and destained with 50% (v/v) acetonitrile and 25 mM ammonium bicarbonate prior to dehydration with 100% (v/v) acetonitrile. Samples were then reduced with 10 mM DTT and 25 mM ammonium bicarbonate for 30 min at RT, followed by alkylation with 20 mM acrylamide and 25 mM ammonium bicarbonate for 30 min at RT. Peptides were extracted from gel pieces with 10% formic acid, followed by 50% acetonitrile/0.1% formic acid and then 70% acetonitrile/0.1% formic acid. Free peptides were recovered from gel pieces and dried in a rotary speed vacuum. Samples were shipped in microcentrifuge tubes for LC/MS/MS analysis.

2.7 LC/MS/MS

Using an Acquity M-class nanoLC system (Waters, USA), 5 µL of the sample was loaded at 15 µL/min for 3 min onto a nanoEase Symmetry C18 trapping column (180 µm x 20 mm) before being washed onto a PicoFrit column (75 µmID x 300 mm; New Objective, Woburn, MA) packed with SP-120-1.7-ODS-BIO resin (1.7µm, Osaka Soda Co, Japan) heated to 45ºC. Peptides were eluted from the column and into the source of a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific, Waltham, MA) using the following program: 5-30% MS buffer B (98% acetonitrile + 0.2% formic acid) over 15 min, 30-80% MS buffer B over 3 min, 80% MS buffer B for 2 min, before column equilibration at 1% for 3 min. The eluting peptides were ionised at 2400 V. A Data Dependant MS/MS (dd-MS^2) analysis was carried out, with a survey scan of 350-1500 Da performed at 70,000 resolution for peptides of charge state 2+ or higher with an automatic gain control (AGC) target of 3e6 and maximum injection time of 50 ms. The top 12 peptides were selected and fragmented in the high energy collision induced dissociation cell using an isolation window of 1.4 m/z, an AGC target of 1e5, and maximum injection time of 100 ms. Fragments were scanned in the Orbitrap analyser at 17,500 resolution and the product ion fragment masses measured over a mass range of 120-2000 Da. The mass of the precursor peptide was then excluded for 30 s.

The MS/MS data files were searched using Peaks Studio X against the UniProt *Rattus Norvegicus* reference proteome database (downloaded May 2022) database and a database of common contaminants with the following parameter settings. Fixed modifications: none. Variable modifications: propionamide, oxidised methionine, deamidated asparagine. Enzyme: semi-trypsin. Number of allowed missed cleavages: 3. Peptide mass tolerance: 10 ppm. MS/MS mass tolerance: 0.05 Da. The results of the search were then filtered to include peptides with a $-\log_{10}P$ score that was determined by the FDR of <1%, the score being that at which decoy database search matches were <1% of the total matches.

3 Results

Although some faint proteoform bands were detected in the extracts from the rehydration trays that had seen face-down rehydration (**Figure 1A**), these proved to be negligible losses overall (<0.25% of the 100 µg loaded onto the IPG strip; **Figure**

1B). Nonetheless, total proteoform losses following face-up rehydration were significantly lower than total losses following face-down rehydration ($t(19) = 7.901$, $p < 0.0001$). Notably, the total signal following face-up rehydration was not significantly different from total signal seen in an empty well extract indicating that there was no discernable loss of proteoforms during face up rehydration using high sensitivity staining and detection protocols (**Figure 1B**). High quality 2D gels (comparable to those seen in previous publications from our group) (**Suppl Figure 1-2**) showed no significant differences in 2D gel spot patterns or counts following face-up (769 ± 54 spots) and face-down (742 ± 42 spots) rehydration (**Figure 1C**).

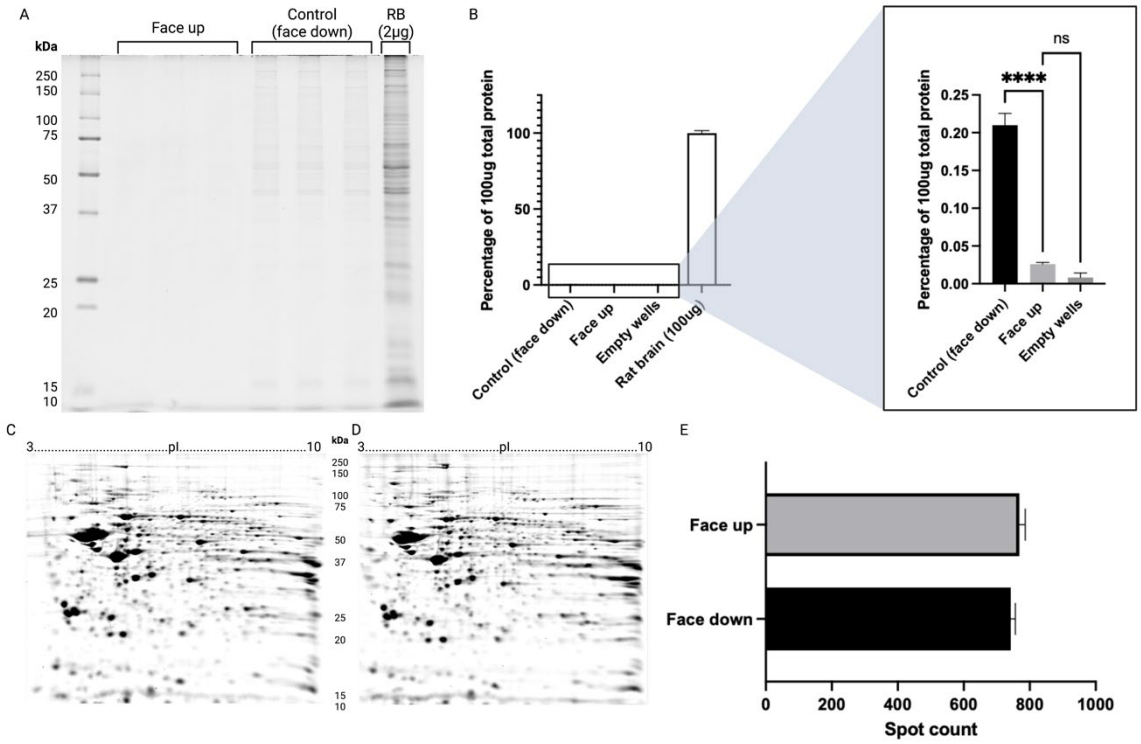

Figure 1. **Proteoform losses to rehydration tray.** (A) Representative 1D gel image of extracts from wells of trays used to rehydrate IPG strips, as well as a sample of total rat brain (RB). (B) Percentage loss compared to 100% load of the RB proteome. Error is SEM. There were n = 3 technical replicates (i.e., lanes) per condition, per experiment. Representative 2D gel image after IPG strips were rehydrated either (C) Face-down or (D) Face-up. (E) Total spot count for 2D gels; n = 9 gels (**Suppl Figures 1-2**). Error is SEM. Experiments were independently replicated 2-3 times. '****' indicates statistically significant difference ($p < 0.0001$); 'ns' indicates no significance.

When analyzing wicks and mineral oil, whether they were unused or had been exposed to IPG strips + 2DB or IPG strips + rat brain extracts in 2DB, a signal was consistently detected at the front of the tris-glycine gels (**Figure 2A**). To better resolve this, tris-tricine gels were used, showing that this prominent band consistently appeared at or near the resolving front (i.e., below ~3.4 kDa) (**Figure 2B**), often as a doublet. Due to this contaminating signal from unused wicks and mineral oil, attempts to analyze possible losses (as described in **Figure 1**) of very small proteoforms/peptides were indeterminate and difficult to interpret (**Figure 2C)**; thus, bands (i.e., below 3.4 kDa in **Figure 2B**) were excised from tris-tricine gels and analyzed by LC/MS/MS to assess for any potential proteoform losses. Notably, there were no Peptide-Spectrum Matches (PSMs) in the samples analyzed by LC/MS/MS with acceptable scores or significance values (**Suppl Figure 4; Suppl Table 1)**. However, even with the contaminating signal, total combined signal from wicks and mineral oil (whether used or unused) was below 0.15% of the 100 μg rat brain proteome extract loaded per IPG strip (**Figure 2C**).

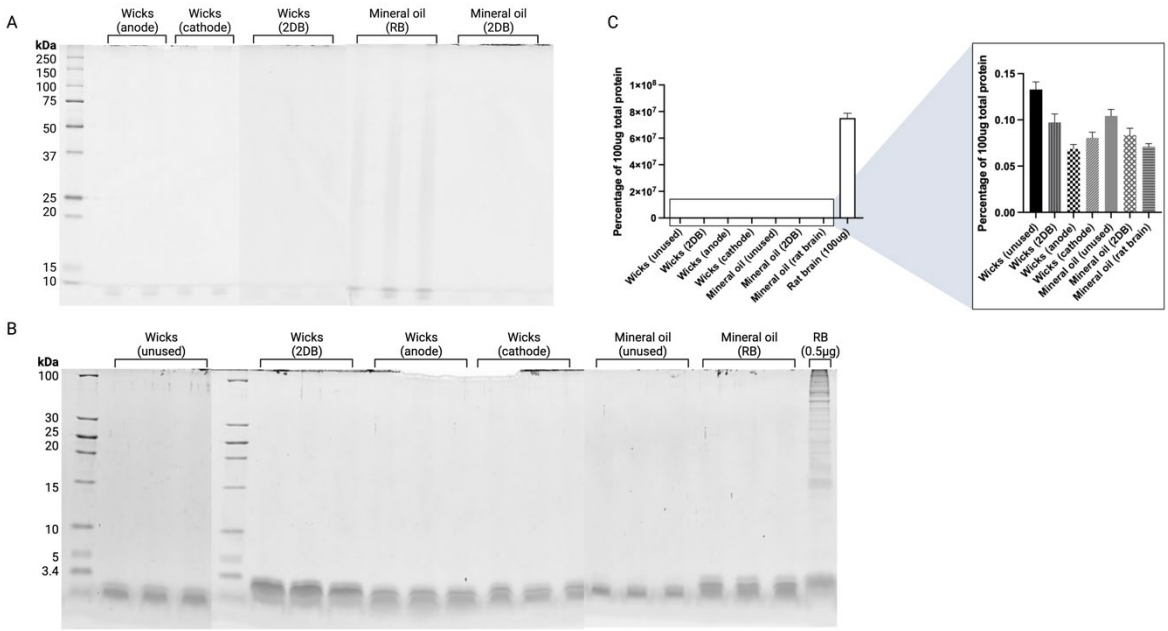

Figure 2. Losses to wicks and mineral oil. (A) Representative 1D tris-glycine gel image of extracts from used and unused wicks and mineral oil as well as total rat brain (RB) sample. (B) Representative 1D tris-tricine gel image of extracts from used and unused wicks and mineral oil as well as a total RB sample. (C) Percent loss compared to 100% load of rat brain proteome. Error is SEM. There were n = 3 technical replicates (i.e., lanes) per experiment. Experiments were independently replicated 2-3 times.

Images in A and B are composites from different gels since sample numbers did not allow for resolution on a single gel.

In order to assess total proteoform recovery and thus quality of subsequent total proteome analyses enabled by 2DE, total signal was compared between 2D gels (face-down and face-up rehydration; **Figures 3A and B**, respectively) and preparative gels (**Figure 3C**), each loaded with 100 μg of total rat brain protein. Notably, there was no significant difference in total protein signals between these three methods of analyzing the rat brain proteome extract (**Figure 3**). Qualitatively, 2D gel spot patterns and intensity were essentially indistinguishable between the two rehydration methods tested.

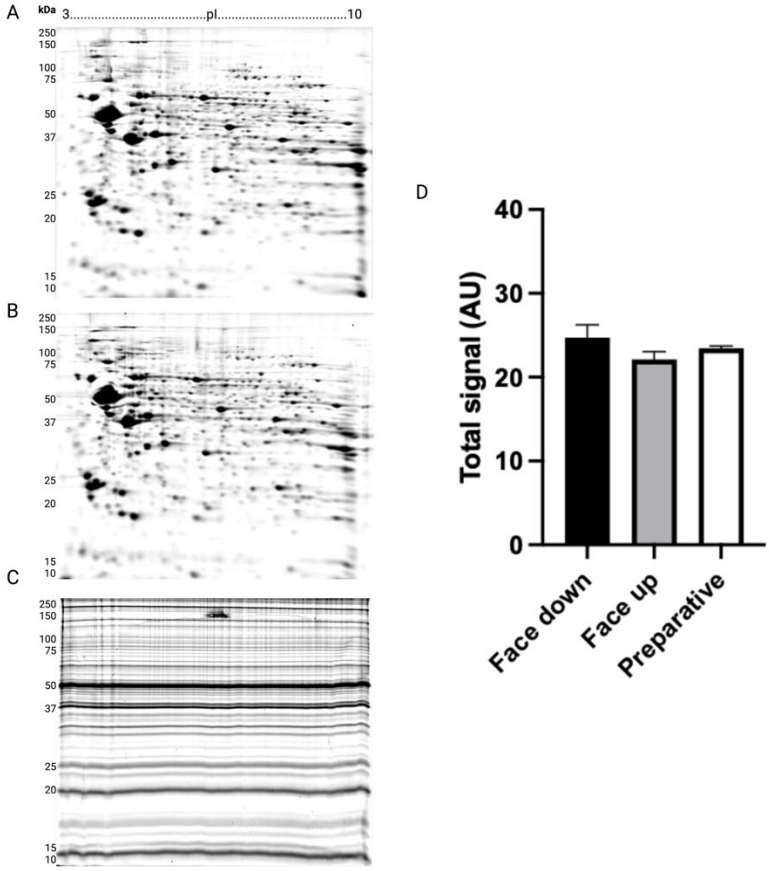

Figure 3. Total protein signal from rat brain proteome (100 μg) resolved by three alternate methods. Comparisons are between 2D gels from IPG strips rehydrated either (A) Face-down or (B) Face-up, and (C) 1D Preparative gels. A-C are representative images of n = 3 technical replicates (i.e., gels) in a single experiment.

(D) Total signals represented by average grey value from images of resolved gels. Error is SEM.

4 Discussion

Much dogma has circulated concerning 2DE, in particular with the introduction of shotgun proteomic analyses yet, as previously noted, much of this consists only of statements repeatedly copied from old review articles [1, 2, 4]. However, the real issue remains: *proteomes are composed of proteoforms*. Thus, any analysis that genuinely seeks to address the actual breadth and depth of proteomes must therefore have the capacity to routinely yield the necessary high-resolution assessments. Currently, only integrative proteomics – high resolution front-end proteoform separation by 2DE and subsequent detailed spot dissection by LC/MS/MS – provides such necessary rigour [3, 4, 24]. In our ongoing efforts to quantitatively test and refine integrative top-down proteomic analyses, here we have evaluated possible proteoform losses during face-down and face-up IPG strip rehydration and the IEF steps of 2DE. In contrast to two ~20-year-old reports of apparent losses during these steps, the data here show that there is no loss of proteoforms to the rehydration tray during face-up rehydration or to the wicks and mineral oil during IEF. Furthermore, while there is a <0.25% loss of total protein to the rehydration tray during face-down rehydration, this does not affect the total spot count or total signal in the final 2D gel pattern. Furthermore, regardless of rehydration method, the patterns and signal from the 2D gels (**Figures 1C and D**) were quantitatively identical. Additionally, the total signals were identical to those from a comparably loaded preparative gel (**Figure 3**), which indicates that there was no discernable loss of proteoforms during the equilibration between the first and second dimensions of resolution.

The results of this study thus differ substantially from two earlier reports [19, 20]. A likely explanation is that noticeable overloading of the IPG strips was being attempted in those earlier studies. In one case, 200 μg of total protein in 160 μL of rehydration buffer was the attempted load on a 7 cm IPG strip [20]. This could well explain the large amount of apparent 'loss' seen considering that the majority of IPG manufacturers and users recommend only loading ~100 μg of total protein in 125-130 μL of rehydration buffer for 7 cm IPG strips [25-27]. Notably, it had previously been reported that overloading IPG strips resulted in an increased loss of protein during the

2DE protocol, and even in this instance the attempt was to load 500 µg of total protein onto 18 cm strips (when recommended loads are ~200-400 µg [19]. Another possible explanation for the different results may be due to the numerous methodological refinements made over the last two decades, including improvements in sample preparation and solubilization, assessing and normalizing total protein concentrations, and standardization of total protein loads per IEF strip (i.e., 100 µg for 7 cm strips and 300 µg for 17-18 cm strips) and possibly rehydration trays made of lower protein binding material than older versions [8, 9, 12].

Nonetheless, fairly prominent bands were routinely detected at the front of tris-glycine gels and below ~3.4 kDa on tris-tricine gels. Notably, these bands were also consistently found in the extracts of unused paper wicks and mineral oil, indicating that they were most likely to be some form of low molecular weight contaminant(s) rather than proteoforms lost from the samples. To test this hypothesis, LC/MS/MS was carried out to determine if any peptides/proteoforms were present in the bands detected below ~3.4 kDa in tris-tricine gels. Notably, there were no positive IDs matching the rat brain database, no evidence of keratin contamination, nor any peptides that could definitively identify a contaminant, suggesting that the contaminant likely consisted of one or more small chemical species. Nevertheless, with cCBB being a highly sensitive nIRFD stain, we conclude that contaminants in the mineral oil and wicks (which had been kept scrupulously sealed and clean, as received from the manufacturers) contributed to the very low MW signals detected, rather than lost proteoforms. Notably, this indicates the importance for future studies to examine if these contaminants affect the results of 2DE and if so, how to remove them from the 2DE protocol.

Although the data here show that face-up rehydration results in no detectable loss of proteoforms to the rehydration tray, wicks, or mineral oil used during IEF, there is an issue with this application method. It was seen occasionally that the rehydration buffer would slip off the IPG strip and go underneath the plastic backing, indicating that this application method should be further improved upon. Until this is complete, we have determined that standard face-down rehydration is an optimal solution as the losses seen to the rehydration tray were negligible.

With 20+ years of refinements, from sample preparation forward – through digestion, LC/MS/MS, and data analysis – including third separations to address the

presence of high abundance species and those at the pH extremes, as well as deep imaging to extract as much data as possible from every gel, the current results further establish that integrative top-down proteomics, using 2DE as the front-end for proteoform separation, provides the best possible breadth and depth of proteome analysis [4, 6-9, 11, 12, 14-16, 28]. Indeed, to understand molecular mechanisms and identify rational therapeutic targets and biomarkers, we must be capable of routinely analyzing these biologically active species rather than only canonical amino acid sequences. We must also be willing to carry out these necessary, deeper analyses. Considering the inherent complexity of proteomes, there is still much to improve in all available methodologies, but integrative top-down analyses clearly provide a rigorous way forward. We thus wish to close by again emphasizing the importance of *continually* assessing and refining methods – not only 2DE, but all analytical methods – and to independently assess previously published methods to ensure quantitative analyses.

Acknowledgements

KC acknowledges Brock University for a Faculty of Graduate Studies Student Bursary. JRC acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC grant number 2019-04324). We thank Jena Yumol, Mike McAlpine, and Wendy Ward for providing rat brains (Brock University ethics certificate no. 18-03-02).

Conflict of interest

The authors have declared no conflict of interest.

Data Availability Statement

The data supporting the findings of this study are available in the supplementary material.

5 References

1. Oliveira BM, Coorssen JR, Martins-de-Souza D. 2DE: the phoenix of proteomics. Journal of Proteomics. 2014;104(C):140-50.
2. Coorssen JR, Yergey A. Proteomics is analytical chemistry: fitness-for-purpose in the application of top-down and bottom-up analyses. Proteomes. 2015;3(4):440-53.
3. Zhan X, Li B, Zhan X, Schlüter H, Jungblut PR, Coorssen JR. Innovating the concept and practice of two-dimensional gel electrophoresis in the analysis of proteomes at the proteoform level. Proteomes. 2019;7(4):36.
4. Carbonara K, Andonovski M, Coorssen JR. Proteomes are of proteoforms: embracing the complexity. Proteomes. 2021;9(3):38.
5. Aebersold R, Agar JN, Amster IJ, Baker MS, Bertozzi CR, Boja ES, Costello CE, Cravatt BF, Fenselau C, Garcia BA, Ge Y, Gunawardena J, Hendrickson RC, Hergenrother PJ, Huber CG, Ivanov AR, Jensen ON, Jewett MC, Kelleher NL, Kiessling LL, Krogan NJ, Larsen M, R. , Loo JA, Ogorzalek Loo R, R. , Lundberg E, Maccoss MJ, Mallick P, Mootha VK, Mrksich M, Muir TW, Patrie SM, Pesavento JJ, Pitteri SJ, Rodriguez H, Saghatelian A, Sandoval W, Schlüter H, Sechi S, Slavoff SA, Smith LM, Snyder MP, Thomas PM, Uhlén M, Van Eyk JE, Vidal M, Walt DR, White FM, Williams ER, Wohlschlager T, Wysocki VH, Yates NA, Young NL, Zhang B. How many human proteoforms are there? Nature Chemical Biology. 2018;14(3):206-14.
6. Coorssen JR, Blank PS, Albertorio F, Bezrukov L, Kolosova I, Backlund PS, Zimmerberg J. Quantitative femto- to attomole immunodetection of regulated secretory vesicle proteins critical to exocytosis. Analytical Biochemistry. 2002;307(1):54-62.
7. Butt RH, Coorssen JR. Postfractionation for enhanced proteomic analyses: routine electrophoretic methods increase the resolution of standard 2D-PAGE. Journal of Proteome Research. 2005;4(3):982-91.
8. Churchward MA, Butt RH, Lang JC, Hsu KK, Coorssen JR. Enhanced detergent extraction for analysis of membrane proteomes by two-dimensional gel electrophoresis. Proteome Science. 2005;3(1):5.
9. Butt RH, Coorssen JR. Pre-extraction sample handling by automated frozen disruption significantly improves subsequent proteomic analyses. Journal of Proteome Research. 2006;5(2):437-48.
10. Butt RH, Coorssen JR. Coomassie blue as a near-infrared fluorescent stain: a systematic comparison with Sypro Ruby for in-gel protein detection. Molecular & Cellular Proteomics. 2013;12(12):3834-50.
11. Gauci VJ, Padula MP, Coorssen JR. Coomassie blue staining for high sensitivity gel-based proteomics. Journal of Proteomics. 2013;90:96-106.
12. Wright EP, Partridge MA, Padula MP, Gauci VJ, Malladi CS, Coorssen JR. Top-down proteomics: enhancing 2D gel electrophoresis from tissue processing to high-sensitivity protein detection. Proteomics. 2014;14(7-8):872-89.
13. Wright EP, Prasad KAG, Padula MP, Coorssen JR. Deep imaging: how much of the proteome does current top-down technology already resolve? PLoS ONE. 2014;9(1):e86058.
14. Epstein JA, Blank PS, Searle BC, Catlin AD, Cologna SM, Olson MT, Backlund PS, Coorssen JR, Yergey AL. ProteinProcessor: a probabilistic analysis using mass accuracy and the MS spectrum. Proteomics. 2016;16(18):2480-90.

15. Noaman N, Abbineni PS, Withers M, Coorssen JR. Coomassie staining provides routine (sub)femtomole in-gel detection of intact proteoforms: expanding opportunities for genuine Top-down Proteomics. Electrophoresis. 2017;38(24):3086-99.
16. Carbonara K, Coorssen JR. A 'green' approach to fixing polyacrylamide gels. Analytical Biochemistry. 2020;605:113853.
17. Bjellqvist B, Ek K, Giorgio Righetti P, Gianazza E, Görg A, Westermeier R, Postel W. Isoelectric focusing in immobilized pH gradients: Principle, methodology and some applications. Journal of Biochemical and Biophysical Methods. 1982;6(4):317-39.
18. Rabilloud T, Valette C, Lawrence JJ. Sample application by in-gel rehydration improves the resolution of two-dimensional electrophoresis with immobilized pH gradients in the first dimension. Electrophoresis. 1994;15(1):1552-8.
19. Zuo X, Speicher DW. Quantitative evaluation of protein recoveries in two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis. 2000;21(14):3035-47.
20. Zhou S, Bailey MJ, Dunn MJ, Preedy VR, Emery PW. A quantitative investigation into the losses of protein at different stages of a two-dimensional gel electrophoresis procedure. Proteomics. 2005;5(11):2739-47.
21. Noaman N, Coorssen JR. Coomassie does it (better): a Robin Hood approach to total protein quantification. Analytical Biochemistry. 2018;556:53-6.
22. Butt RH, Pfeifer TA, Delaney A, Grigliatti TA, Tetzlaff WG, Coorssen JR. Enabling coupled quantitative genomics and proteomics analyses from rat spinal cord samples. Molecular & Cellular Proteomics. 2007;6(9):1574-88.
23. Schagger H. Tricine-SDS-PAGE. Nature Protocols. 2006;1:16-22.
24. Naryzhny SN, Zorina ES, Kopylov AT, Zgoda VG, Kleyst OA, Archakov AI. Next steps in silico 2DE analyses of chromosome 18 proteoforms. Journal of Proteome Research. 2018;17(12):4085-96.
25. IPGphor Isoelectric Focusing System: User Manual. 2000 [cited 2022]; Available from: https://www.bu.edu/picf/files/2010/10/IPGphorManual.pdf.
26. Instruction Manual: SERVA IPG BlueStrips. 2022; Available from: https://www.serva.de/www_root/documents/Manual%20SERVA%20IPG%20BlueStrips_43001_25_e.pdf.
27. ReadyStrip IPG Strip Instruction Manual. [Instruction Manual] 2022; Available from: https://www.bio-rad.com/webroot/web/pdf/lsr/literature/4006166G.pdf.
28. Carbonara K, Coorssen JR. Sometimes faster can be better: microneedling IPG strips enables higher throughput for integrative top-down proteomics. Proteomics. 2022:e2200307.

Supporting information

Supporting information file:

Word document - Suppl Figure 1: 2D gels following face-down rehydration; Suppl Figure 2: 2D gels following face-up rehydration; Suppl Figure 3: 1D preparative gels; Suppl Figure 4: MS/MS spectrum of peptide sequence AAHPKLFLL.

Excel document – Suppl Table 1: Reported Peptide to Spectrum Matches.