

Proteomics—The State of the Field: The Definition and Analysis of Proteomes Should Be Based in Reality, Not Convenience

Jens R. Coorsen ^{1,2,*}  and Matthew P. Padula ^{3,*} 

¹ Department of Biological Sciences, Faculty of Mathematics and Science, Brock University, St. Catharines, ON L2S 3A1, Canada

² Institute for Globally Distributed Open Research and Education (IGDORE), St. Catharines, ON L2N 4X2, Canada

³ School of Life Sciences and Proteomics, Lipidomics and Metabolomics Core Facility, Faculty of Science, University of Technology Sydney, Sydney, NSW 2007, Australia

* Correspondence: jrcoorssen@gmail.com (J.R.C.); matthew.padula@uts.edu.au (M.P.P.)

Abstract: With growing recognition and acknowledgement of the genuine complexity of proteomes, we are finally entering the post-proteogenomic era. Routine assessment of proteomes as inferred correlates of gene sequences (i.e., canonical ‘proteins’) cannot provide the necessary critical analysis of systems-level biology that is needed to understand underlying molecular mechanisms and pathways or identify the most selective biomarkers and therapeutic targets. These critical requirements demand the analysis of proteomes at the level of proteoforms/protein species, the actual active molecular players. Currently, only highly refined integrated or integrative top-down proteomics (iTDP) enables the analytical depth necessary to provide routine, comprehensive, and quantitative proteome assessments across the widest range of proteoforms inherent to native systems. Here we provide a broad perspective of the field, taking in historical and current realities, to establish a more balanced understanding of where the field has come from (in particular during the ten years since Proteomes was launched), current issues, and how things likely need to proceed if necessary deep proteome analyses are to succeed. We base this in our firm belief that the best proteomic analyses reflect, as closely as possible, the native sample at the moment of sampling. We also seek to emphasise that this and future analytical approaches are likely best based on the broad recognition and exploitation of the complementarity of currently successful approaches. This also emphasises the need to continuously evaluate and further optimize established approaches, to avoid complacency in thinking and expectations but also to promote the critical and careful development and introduction of new approaches, most notably those that address proteoforms. Above all, we wish to emphasise that a rigorous focus on analytical quality must override current thinking that largely values analytical speed; the latter would certainly be nice, if only proteoforms could thus be effectively, routinely, and quantitatively assessed. Alas, proteomes are composed of proteoforms, not molecular species that can be amplified or that directly mirror genes (i.e., ‘canonical’). The problem is hard, and we must accept and address it as such, but the payoff in playing this longer game of rigorous deep proteome analyses is the promise of far more selective biomarkers, drug targets, and truly personalised or even individualised medicine.

Keywords: integrative top-down proteomics; bottom-up proteomics; immunoassay; mass spectrometry; protein species; proteoforms; structural prediction; tandem MS; two-dimensional gel electrophoresis; Western blotting; multi-omics; isoelectric focussing



Citation: Coorsen, J.R.; Padula, M.P. Proteomics—The State of the Field: The Definition and Analysis of Proteomes Should Be Based in Reality, Not Convenience. *Proteomes* **2024**, *12*, 14. <https://doi.org/10.3390/proteomes12020014>

Academic Editor: Jörg Reinders

Received: 17 March 2024

Revised: 17 April 2024

Accepted: 17 April 2024

Published: 19 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abbreviations

2DE	Two-dimensional gel electrophoresis
IEF	Isoelectric focusing
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
MW	Molecular weight

MS	Mass spectrometry
LC	Liquid chromatography
TMS	Tandem mass spectrometry
TDP	Top-down proteomics
iTDP	Integrative top-down proteomics
MSi-TDP	Mass spectrometry intensive top-down proteomics
BU	Bottom up
BUP	Bottom-up proteomics
ORF	Open reading frame
PTM	Post translational modification
FTICR-MS	Fourier transform ion cyclotron resonance mass spectrometry
LFQ	Label-free quantification
pI	Isoelectric point

“We choose to go to the Moon in this decade and do the other things, not because they are easy, *but because they are hard.*” John F. Kennedy; address at Rice University, Houston, Texas, 12 September 1962 (our italics).

1. Introduction

With origins most logically traced to the development of two-dimensional gel electrophoresis ((2DE) combining isoelectric focusing (IEF) and SDS-PAGE in progressive dimensions of separation) which enabled the resolution of hundreds (likely many thousands) of proteoforms/protein species [1–10], proteomics has undergone notable changes. Along with other versions of gel-based 2D separations (i.e., chromatography in a gel matrix), changes have included almost five decades of 2DE optimization yielding a truly high-resolution analytical platform for proteoforms across very broad ranges of charge (pI) and molecular weight (MW). Despite the development of new technologies in the intervening decades, *no other methods* provide such capacity for genuinely deep, comprehensive proteome analysis at the essential level of proteoforms [7,10,11]. With the introduction of mass spectrometry (MS) and particularly its coupling to liquid chromatography (LC), the coupling with tandem MS (LC/TMS) proving most productive at scale [12–15], proteomics has become a discipline in its own right. This combination of high-resolution technologies—2DE/LC/TMS—fully enables the highest resolution analytical chemistry approach to proteome analysis that is now most widely and appropriately referred to as top-down proteomics (TDP) [7–10,16,17].

Piggybacking on developing gene and amino acid sequencing methods, and the then pending and subsequent first release of the human genome, an alternate—bottom-up (BU) or ‘shotgun’—approach to proteome analysis came into vogue [18–20]. Notably, in the 1990’s, gene and protein sequencing faced similar problems in that methods were slow and increasingly unreliable with larger molecular sizes. For genomics, this led to short-read DNA sequencing in which 100–200 ‘reads’ ensure accuracy. BU proteomics (BUP) has sought increased speed of canonical protein identifications (i.e., linkage to recognized/canonical gene sequences) rather than analytical depth at the critical level of proteoforms. Thus, this purely proteogenomic approach relies on gross digestion of complex native proteome extracts, LC for rudimentary sorting of the resulting peptide milieu (i.e., seeking to reduce the resulting complexity), TMS to sequence individual peptides, and software to link these to protein sequences or predicted Open Reading Frames (ORFs) in databases. Much of this is said to be undertaken in a relatively unsupervised manner, and purportedly with less hands-on technique relative to 2DE/TMS approaches. Notably, the replicate determinations recognised as critical in DNA sequencing never seem to have been viewed with the same rigour by most BUP practitioners (but are fortunately standard practice in most 2DE/TMS analyses) [7,8,11]. Shortcomings and the need for more routine rigour in BUP have only become slowly apparent, as is not unusual with new research approaches [7,21–30]. Nonetheless, while initially providing some ease in terms of analytical methodology, and reasonable coverage of high abundance canonical amino acid sequences, the vast majority of BUP identifications are inferences based often

on only one or a couple of peptides (i.e., not even approximating the full length of the amino acid sequence in databases). Highly abundant sequences tend to dominate the findings. However, beyond the widely recognized problems of inferring canonical protein identifications and assuming the presence of intact/full length species, the most serious issue by far is the loss of all information concerning *specific* proteoforms. Indeed, early work with this approach even ‘corrected’ datasets to eliminate isoforms and PTM in order to simplify database searching and focus only on canonical amino acid sequences. This is another example of supposed simplification obscuring rather than addressing proteome complexity. Unfortunately, recent BUP representations of proteoform ‘groups’ derived from the identification of modified peptides also do not address the critical issue of quantitatively identifying *specific* proteoforms [31]. This merely extends the inference problem by vaguely indicating that *some* proteoforms are apparently present (or at least a modified peptide is part of the peptidome). Essentially, this only moves us from an inferred ‘canonome’ to the speciation of peptides. In the end, such inference, rather than definitive resolution and identification, is comparable to epidemiology in that it cannot establish definitive characteristics or causes (i.e., a specific proteoform) but only correlations (i.e., of a peptide to a sequence in a database). Nonetheless, should a case arise in which a specific proteoform is, for example, definitely linked to a function or disorder, then a targeted assessment of one or more of its distinctly characteristic (i.e., modified) peptides might serve for screening.

Therefore, we submit that BUP as currently and popularly applied is a quick cataloguing tool but, in and of itself, cannot provide the data needed for deep, critical, systems-level understanding of biological processes [10,32–36]. Identification pipelines that reject certain ‘protein’ identifications because, a decade or more ago, they were thought to be ‘contaminants’ or ‘routinely appearing’ [37] need to have their appropriateness reassessed. As proteoforms were never identified in the original studies using BUP identification approaches, one must consider that these studies and resulting databases ignore potentially critical proteoforms as the original analyses considered only canonical amino acid sequences correlated years ago with gene databases. It is, however, crucial to note that BUP is a powerful component of Integrated or Integrative TDP (iTDP) analytical approaches (see below).

2. Where Things Stand and Why

Realistically, we are now in the *post-proteogenomic era*, and likely (should) have been for well over a decade or more already. Change can be hard but only genuinely deep, comprehensive proteome analyses at the critical level of proteoforms will provide the data necessary to identify rational biomarkers and therapeutic targets via extensive dissection of molecular mechanisms and pathways. This critical approach will provide objective, systems-level understanding of biology, in particular coupled with a growing appreciation of genome complexity, gene regulation, epigenetics, and metabolomics. Indeed, a recent suggestion is that the field should be doing “proteoformics”—focusing on proteoforms—rather than proteomics [38]. The reader can draw the parallel to the JFK quote above, that researchers in proteomics need to focus on what is hard because it will ultimately provide the most critical, informative, and thus essential data. Continuing on the current ‘easy’ path brings to mind Einstein’s (misattributed) quote that “Insanity is doing the same thing over and over and expecting different results” [39].

In this Perspective, we will take a hard look and ask some hard questions as it is high time for the field to come to grips with them. To understand how we intend to identify/address these issues, it might be best for readers to first consider the ‘what if’ questions relative to what is now finally acknowledged as the real complexity of proteomes [10,33–36,40–44]. What if the field had progressed differently? Think deeply and purely objectively for a moment about what has transpired and in terms of the present and future. What if the field had not taken a predominantly proteogenomic strategy over the last 20+ years? Minimally, we knew at the time that there were undoubtedly a substantial number of variants to any gene product or protein (e.g., mutations, alternate and multiple

reading frames, splice variants, posttranslational modifications (PTM) [1,45]), and thus that BUP could never provide the depth of analysis needed for genuinely comprehensive proteome analyses [46–48]. Nonetheless, ‘fast’ and ‘easy’ analyses linking only to the genome became the standard. Some prominent journals even established BUP as their sole focus, rejecting any other studies, even when they employed higher resolution approaches. Notably, it was also already known that 2DE coupled with MS could resolve and identify many thousands of proteoforms from proteome extracts [2,49–54]. What if, rather than going down the purely proteogenomic rabbit hole, we had fully utilised the truly high resolution 2DE/LC/MS analyses—a *genuine, comprehensive and integrative TDP approach* [3–6,8,10,17,55]—how many (human) proteoforms would our databases already contain, minimally at the level of pI and MW if not actual PTM, along with the canonical amino acid sequence? How many more selective antibodies (and antibody therapeutics) would already be available? How many highly specific/selective biomarkers and drugs would already be available (or at least close to market) to address critical healthcare burdens? Even individualised medicine? We leave it to the concerned reader to provide a critical if only conservative estimate. A better, transparent, more collegial, complementary, and thus integrated path forward is clearly needed.

3. What Is Proteomics? What Is a Proteome? Defining Issues to Date

The first issue is recognition that proteomics deals exclusively in proteoform abundance, *not* protein expression or up/down-regulation. The latter require different assays. Even when correlations with mRNA levels happen to exist, these do not fully establish that changes in canonical protein levels are due solely to changes in gene expression unless stability of the species (e.g., degradation rate) is also assessed. We consistently see such terms misused in the literature, particularly in studies not published in rigorous proteomics journals. In this regard, it is also frustrating to see references to a gene having or doing a certain function; genes are codes, not functional/active entities, and thus do not otherwise ‘do’ anything.

Appropriately defining a proteome is the next and more important issue that must be addressed in developing a truly comprehensive (i.e., deep and quantitative) and broadly unified analytical approach. Although the term “proteome” was first coined by Marc Wilkins and colleagues in 1995 [56], realisation of the genuine complexity of proteomes now demands that, beyond simply the canonical amino acid sequences encoded by the genome, proteomes be most accurately defined by their proteoform constituents. This should then be further refined by location/space (e.g., particular cells, subcellular compartments/organelles, molecular complexes) and time, as the proteome is highly dynamic relative to the genome and transcriptome. Despite these critical considerations, to ensure the highest quality proteome analyses, a core issue remains the criteria for defining a proteome.

Regrettably, in considering the literature over the last decade or more, ‘proteome’ seems largely to have become a term of convenience rather than rigour. To fully address the complexity of proteomes, the simplest first step is recognition that the word ‘proteoform’ should most appropriately replace the generic term ‘protein’ in almost all usage other than general references to that latter group of macromolecules; in the case of proteogenomic/BUP studies, the data are most appropriately referred to as ‘inferred ORF products’ [45]. Perhaps it would help if that change was made in Wikipedia so that, earlier in their education, students already understand and accept the real complexity inherent to proteomes? Simply continuing to introduce students to the Central Dogma in high school is painfully outdated and insufficient for their future endeavours, or for research efforts overall.

However, rather than such a critical and objective approach, defining the proteome seems to have become a method-dependent matter of convenience. For BUP studies, this largely means inferred identification and quantification of apparent proteogenomic/canonical ORF products without mention of the lack of isoform or proteoform identifications or the problems of protein inference. For MS-intensive TDP (MSi-TDP) [7,57,58], this essentially

means some sub-proteome, largely if not solely within the <20–30 kDa size range of total species in the proteome [59]; while these methods are very occasionally able to analyse a higher MW species, this amounts to a vanishingly low fraction of any given proteome [60]. Although some published studies refer to these as ‘comprehensive’ proteome analyses, this is at best misleading. While MSi-TDP can provide incredibly detailed assessments of proteoforms and isolated complexes that conform to the capabilities of the method, these are not quantitatively or even qualitatively deep, comprehensive assessments of proteomes (i.e., across a breadth of pI and MW that defines native proteoforms in cells, tissues, and biological fluids). Is this, and the very expensive, high-end MS instrumentation required (e.g., FTICR MS), a promising approach? Absolutely. But, there have been limited advances in this approach for the last 1–2 decades due largely to the diversity and dynamic range of the proteome (e.g., notably mid-to-large MW and membrane species), poor front-end resolution of species even when combining multiple separation steps (i.e., still resulting in co-elution of small and large species), the decay in signal-to-noise with large species due to a plethora of charge states, and the need for better software integration, as well as even more effective dissociation methods to fragment larger species [58,61–64]. Indeed, while smaller proteins have been analysed by MS since the early 1960s, including with early software programs [65–70], some of the first reports of mid-to-large proteins being analysed by MSi-TDP were from the early 1990s until the mid-2000s [58,71–74]. Thus, there has been little substantive advance in terms of a broad and consistent breaking of the current ~20–30 kDa ‘MW barrier’. Routine full proteome analyses using the MSi-TDP approach must still await further developments and rigorous testing. Perhaps its strongest immediate application is in the analysis of the small proteoforms inherent to isolated proteins and protein complexes [75–80]. In this regard, it is also important to note recent advances in MSi-TDP instrumentation that enable effective analysis of low MW species by capitalising on the practical mass resolution of instruments used for BUP, these being able to resolve the isotopic series of proteoform charge state (e.g., Exploris 120/240/480, TIMSTOFs, ZenoTOFs) (e.g., [81]). Accordingly, MSi-TDP can and should be used by BUP practitioners as a complementary technique enabling a more critical analysis of at least a fraction of the proteome. Furthermore, considering inherent issues with the front-end protein separation methods routinely used in MSi-TDP—e.g., the misleadingly acronymed Gel-Eluted Liquid Fraction Entrapment Electrophoresis (GELFrEE), a 1D separation utilising tube gels and SDS which must be removed prior to LC or MS [82]—employing 2DE, which, by design, ‘isolates’ proteoforms during the two-step resolving process, would likely promote deeper analyses, perhaps even of larger species. To date, such a critical approach remains untested [83]. MSi-TDP practitioners continue to use combinations of GELFrEE and/or multiple LC phases, despite recognised issues of co-eluting large and small species and complex spectra that require multiple software tools for downstream analyses that can take multiple hours or even longer to complete yet can still yield ambiguous identifications.

Thus, rather than methods-centric, methods-dependent, or otherwise insular definitions, the simplest, most objective and straightforward definition of a proteome is that it is a specific collection of proteoforms that are intrinsic to the native state at the time of sampling. What that native state is must be specifically defined: sample source/type, including specific (sub)fractions; all details of sample handling/processing and any ‘fractionation’; all details of downstream analysis including specifics of any and all resolving protocols (e.g., gel and/or LC, and MS), data processing, and final broad availability of the data. The latter has, of late, become an interesting concern as genome and protein analyses and databases become more refined, leading to the question of how many canonical protein identifications (e.g., from a decade or more ago) would still be accurate if reassessed? Perhaps there are already critical data in the literature that have been ‘missed’ and, likewise, red herrings misleadingly identified as critical that should not have been a focus had better data/interpretation been the original outcome. Therein lies the critical need for more complete amino acid sequence coverage and resolution of proteoforms [84–89], constant refinement of statistical and bioinformatics tools, and data banking [10]. Furthermore, raw gel images, LC chromatograms, and MS data for all published studies must be made

available on publicly accessible data repositories. While this is standard rigour at all critical proteomics journals, many medical and other journals still do not demand this assurance of data reproducibility, often enabling the publication of proteomic research of questionable quality. In summary, any study referred to as unbiased, global, (ultra)deep or otherwise ‘comprehensive’ when it does not address proteoforms but only canonical amino acid sequences and/or ‘sub-proteomes’ within only limited MW ranges, or that does not fully establish its reproducibility, is quite unrealistic and disingenuous.

Fundamentally, it is only the ‘detectable’ proteome that is defined by the methodologies employed. To address limitations of methods/instrumentation, it has become common practice to reduce sample complexity, with all the inherent risks that entails. To reduce data complexity and aid ease of interpretation, the sample is fractionated according to physicochemical properties (e.g., proteoform size and/or charge, peptide hydrophobicity, complex size), either to focus on a compartment (e.g., nucleus, membrane, vesicle) or increase identification depth (qualitatively, sometimes quantitatively). In this vein, PTM characterisation using BUP currently favours enrichment of the modified peptides in an effort to overcome their generally lower abundance in the total peptide milieu compared to the unmodified peptides. Such approaches not only disconnect PTM from the specific proteoform (see above) but also bias research toward PTM that have purportedly reasonable enrichment strategies. This is why phosphorylation is the most studied PTM. Therefore, how do multi-step fractionation and enrichment protocols affect sample quality (e.g., proteoform/PTM lability) and thus the qualitative and quantitative accuracy of analyses? Again, the concept of proteoform groups, based on peptide speciation in BUP, is a methods-limited extension of the inherent limitations of inference. In the end, is there really any 100% suitable solution to truly comprehensive proteome analysis aside from analysing the native sample as close to its native state as possible?

4. Recognising and Addressing Critical Issues

To begin, we should highlight progress made in the last ~30 years, using the first report to correlate peptide data with canonical amino acid sequences in databases [18] as a benchmark, and progress made in the 10 years of the journal *Proteomes* existence. There is certainly no debate that inference of canonical ‘protein’ identities based on peptide data from a shotgun analysis has become the most widespread approach in proteomics. Although BUP has been most concerned with a suggested speed or ease of analysis and correlations with canonical gene sequences (i.e., proteogenomics) [90], it has to some extent also sought to be quantitative where possible. Thus, label-free quantification (LFQ), while clearly having some notable limitations [91–94], including the inherent failure to discern peptide distributions between proteoforms [32,95] and the missing values problem [96–99], can prove reasonably informative when used judiciously [48]. Similarly, a range of peptide labelling techniques (e.g., Tandem Mass Tags) have sought to routinely quantify changes in the abundance of canonical proteins, although there are quite critical concerns to be taken into account in employing any labelling methods [10,100,101]. What then are some of the most critical concerns and advances in proteome analysis to date and, respectively, how must these be addressed and further developed and optimised to ensure genuinely deep, quantitative proteome analysis at the level of proteoforms?

4.1. Improvements in Proteoform Extraction and Sample Processing

Numerous concerns arise from the outset of sampling [102–104]. We base this on our firm belief (or mantra) that *the best proteomic analyses reflect, as closely as possible, the native sample at the moment of sampling*. Although desirable to immediately process the sample for analysis, this is clearly not feasible in most studies. Some studies take care to sample and store as quickly as possible (e.g., immediately snap freezing in liquid nitrogen and storing at $-80\text{ }^{\circ}\text{C}$), while others likely incur artefacts by prolonged handling or further processing at room temperature (or higher), either pre- or post-freezing, and/or simply slow freeze the sample by placing it directly at $-30\text{ }^{\circ}\text{C}$ or $-80\text{ }^{\circ}\text{C}$. While this problem may

well be sensitive to sample volume, it appears to concern most research on blood serum and plasma. These are, in fact, far more problematic in terms of sampling, which will vary according to the gauge of needle used during phlebotomy, the length of time and the temperature at which the whole blood samples are kept prior to and during processing, and which anticoagulant is used to collect plasma or how the blood is allowed to coagulate to collect serum. Taking all these critical factors into account, it is clear that, in many studies, inter- and intra-sample variability mean that neither serum nor plasma represent 'native' or well-controlled replicates [105–108]. Considering cell and platelet lysis or activation that occurs during phlebotomy and/or blood processing, one must seriously ask about the definitions of serum and plasma—are these meant to represent components free in the blood during circulation *in vivo* or to refer to the total complement of components in the blood, including the contents of cells and platelets? It has also become clear that what is lost to the red cell fraction must also be taken into account [109,110]. Again, from a proteomics (or any 'omics) perspective, it is critical to know all details of sampling and processing. This becomes particularly relevant in searching for biomarkers or potential therapeutic targets.

Regarding sample extraction, testing new detergents, chaotropes, and surfactant combinations has been a critical focus in proteomics since the development of 2DE [111–113]. The challenge both for the first dimension of 2DE and for MS analyses has been the incompatibility of SDS, which is broadly considered the gold standard for proteoform extraction. Nonetheless, effective combinations of automated frozen disruption, detergents, and denaturing agents have enabled even the resolution of complete membrane proteomes by 2DE, despite long-held dogma suggesting this was not feasible [4,113–121]. Indeed, refined extraction and 2DE protocols have established quantitative proteome analysis consistent with SDS extraction [122]. Cleavable (e.g., acid-labile) or photodegradable surfactants have also been used in BUP and MSi-TDP but these do not appear to have been widely adopted despite extraction potentially comparable to SDS [123,124]; these may negatively impact PTM and do not help overcome the MW limitations of MSi-TDP analyses. Another new avenue includes the use of ionic liquids and other extraction media to improve recovery of otherwise insoluble proteoforms. In the case of ionic liquids, there is evidence that certain combinations cause artefactual modification of proteoforms through backbone cleavage or side chain modification [125], requiring further investigation of these otherwise promising extraction reagents. Overall, depending on the sample type and/or focus of the analysis, it would be prudent to ensure optimization of extraction conditions, in particular if total proteome analysis is the goal. It is unlikely that one size fits all when it comes to total extraction of proteoforms (or as close to it as is possible).

Evidence that sample reduction (to effectively remove disulfide bonds) has been largely underpowered in most studies to date, and that optimization of this critical step further enhances TDP analyses, again emphasises the need to continuously and rigorously evaluate even long-established protocols [126]. The caveat remains, however, that we can never be fully certain that we have quantitatively recovered every copy of every possible proteoform from any given sample or sample type, particularly when recovery steps (i.e., centrifugation) are used to remove notable 'insoluble' materials, which seems more often to represent less quantitatively rigorous methodology. This is also particularly true when any purification steps are used in an attempt to isolate specific cellular fractions or proteoforms [127]. In affinity or other fractionation methods, appropriately thorough analyses demand that both resulting fractions be analysed in order to establish and account for the quantitative capacity of the protocol [128]. This is only quite rarely seen in the proteomics literature but is most consistent with good analytical practice.

While we unfortunately continue to see studies that still fail to use any or only minimal protease inhibitors during sampling, the use of broad spectrum protease, kinase, and phosphatase inhibitors, while not an exhaustive approach, was introduced two decades ago in an effort to preserve the native state of proteoforms as best possible [114,129]. This practice should be extended to include other proven small molecule inhibitors of PTM reactions; notably, however, this will also increase the cost of analyses. In the long-term, though,

can we afford to get this wrong in terms of identifying critical proteoforms? Similarly, many studies utilise processing protocols in which proteome extracts (or peptides) are maintained at room temperature or higher for an hour or more (e.g., during chemical labelling protocols) [130]. It seems quite unlikely that such samples effectively reflect the native state of the proteome at the time of sampling, particularly in terms of labile PTM. We know for example that extended incubations even at reduced temperatures result in proteome/proteoform alterations [131]. Overall, then, sample processing times and conditions are of concern. Focused analysis of this issue would be useful to characterise and quantify the extent of proteoform changes, in particular to PTM. Furthermore, consistency and sensitivity in assessing total protein concentration in samples is essential to ensure quantitative comparisons [132]. Accordingly, the practice of assaying before organic precipitation, for example, but assuming total and consistent recovery of species should not be continued. Normalisation must be to the total protein content of each sample to be analysed.

4.2. Improvements in Proteoform Resolution by 2DE

Relative to LC and MS, 2DE protocols and instrumentation have likely undergone comparable if not greater refinement and optimization over the last 30 years; this now enables the deepest proteome analyses currently available by providing high resolution separation of the proteoforms inherent to any sample [4,5,10,55,122]. Indeed, deep analyses have confirmed that resolved ‘spots’ on 2D gels—the macro ‘visible’ overlapping of staining signals from many resolved micro spots of separate species migrating to almost the same location—contain multiple proteoforms. This enables reasonable estimations that a refined and optimised iTDP approach can resolve and identify ≥ 1 M proteoforms across a large pI and MW range, including low abundance species [4,55]. *There is no other current or developing analytical approach that provides such routine or deep proteome analyses.* In contrast to BUP, it is also important to note that a single spot from a 2D gel is a relatively simple sample compared to a gross, whole proteome digest, and likely the reason for much higher routine sequence coverage of species by iTDP.

The commercial availability of quality-controlled isolated pH gradient (IPG) gel strips for IEF established a consistent and thus highly reproducible first dimension for proteoform resolution [133]. Some might suggest the same is true of commercially available SDS-PAGE gels for the second dimension, although consistency in self-casting is easily achievable (e.g., using multi-casting chambers), and this also enables critical fine-tuning of gel composition (i.e., % acrylamide or all important gradient gels) and detergent choice/combination to optimise proteoform resolution depending on the nature of the sample [114,128]. Regrettably, this is often overlooked in favour of the convenience of precast gels, despite their cost, limits of resolution, and resulting plastic waste from the cassettes.

Considering the unparalleled resolving power of 2DE, it is quite disappointing to find that many studies using this technique fail to report the pI of species of interest. This is critical to fully capitalise on the resolving capacity of highly refined 2DE protocols by calibrating both the first and second dimensions of separation using appropriate standards and reporting both the pI and MW of species analysed. Relative to the canonical values in databases (i.e., calculated purely based on the amino acid backbone), this is the most straightforward first confirmation that a species of interest is a specific proteoform [4,116,117,119–121,128,134–137]. Staining 2D gels with PTM-selective stains (e.g., phospho- and glyco-protein reagents) can also provide front-end confirmation of certain modifications and, considering that these reagents can often be used in conjunction with total proteoform detection (e.g., colloidal Coomassie Brilliant Blue (cCBB)), it behoves researchers to extract as much information per gel as possible [4,128,138,139]. Well-planned studies can also capitalise on third separations (i.e., 2DE/3DE) to further resolve species obscured by hyper-abundant spots, as well as those at pI extremes and the gel front, providing still deeper proteome coverage within a single experiment, and further ‘simplifying’ subsequent MS analyses [4,114,128].

Of most recent note, the overall 2DE process has also undergone a substantial increase in throughput. Micro-perforating (i.e., microneedling) IPG strips significantly reduces rehydration loading time, thus saving about a day in the overall 2DE process and establishing that 'faster' processes can also be 'better' without sacrificing the quality or depth of quantitative analyses [140]. While there have also been reports of faster second dimension separations (i.e., PAGE), these rely on higher voltages and (local) temperatures and thus likely result in proteoform artefacts. Resolving the first and second dimensions at lower temperatures would still appear to be the best available approach [4,114,141]. Subsequent to 2DE, an alternate gel fixation protocol, avoiding organic solvents, further minimises the likelihood of artefactual alterations to resolved proteoforms [142]. Thereafter, numerous stains are available to detect the total proteome [138,139]. Of these, the most significant recent innovation may well be the development of a high sensitivity (e.g., femto-to-attomole) detection protocol that uses cCBB as a near-infrared dye [143–146]. Together with the development of a deep-imaging protocol to overcome the signal saturating effects of highly abundant species, this combined optimised approach can detect and identify even low abundance proteoforms within the total proteome [147]. In part, this is also due to the development and commercial availability of (1) high-resolution imaging equipment supporting multiwavelength excitation and emission (e.g., Odyssey imager (Licor, Lincoln NB); Typhoon™ FLA-9000 (GE Healthcare); Typhoon 5 Biomolecular Imager (GE Healthcare)); and (2) image analysis software enabling high resolution spot identification (i.e., signal above local background) and detailed quantitative analyses (e.g., Delta2D (DECODON)). There is also ongoing development of quantitative image analysis approaches that may lead to the critical extraction of still more/better data from 2D gels [148–153]. Furthermore, the immunoblotting of 2D gels, even after staining, provides the most direct approach to quickly identifying proteoforms [3,52,154–165], provided the antibodies used have been critically vetted (including to PTM [166]), and even then there are caveats to consider (e.g., blockade of antibody binding by a given PTM) [10]. Furthermore, a well-established immunoblotting method ensures detection sensitivities in the femto-to-attomole range [141,167]. Often ignored, however, is the need to consistently assess transfer efficiency in order to ensure truly quantitative assessments; this is the analytical equivalent of assessing both the eluate and the retentate in rigorous affinity analyses.

At the interface of BUP and iTDP approaches is the need to digest samples for subsequent peptide analysis. One of the notable advances in this area has been the recognition that strong digestion conditions (i.e., high protease concentrations and/or the standard 37 °C) result in loss of lower abundance species and/or the overwhelming of their MS signals by protease autolysis products. Reducing both protease concentrations and incubation temperature yields far more reliable data in the form of increased sequence coverage per species assessed [4,120,121,137,147,168–170]. Additionally, there have been several reports of (ultra)fast digestion approaches although none seem to have come into widespread use, perhaps raising questions of quantitative losses [169–173].

4.3. Improvements in Liquid Chromatography

Regardless of the improvements in the speed and sensitivity of mass spectrometers, it is still impossible to analyse a whole proteome without some form of fractionation [174,175]. In BUP, many peptides will be essentially isobaric in their m/z value, leading to co-isolation for fragmentation, and thus complicated MS/MS spectra containing fragments from multiple peptides from different proteoforms [176–178]. In electrospray ionisation (ESI), it is well established that samples that are too complex cause ion suppression and non-detection of some analytes [179]. Thus, LC will continue to be an essential part of proteome analysis, whether coupled to MS or not. The current focus of BUP practitioners is increased throughput to obtain large datasets with comparable sample numbers to next generation genomic sequencing (NGS), to increase statistical power if not actual deep proteome coverage [29,180,181].

Despite assertions by instrument manufacturers' marketing departments, LC as used in proteomics is often the last realm of the tinkerer, with set ups varying greatly from platform to platform depending on the operator. Maximum sensitivity, especially for analysis of 2DE-resolved proteoforms and single cell analysis, still requires nanoflow chromatography (<1 uL/min) using in-house manufactured columns with integrated ESI emitters that can reveal column-to-column reproducibility issues if these are not carefully quality-controlled [182]. A vendor-promoted move to microflow LC (1–5 uL/min) has resulted in more robust and reproducible separations and peptide ionisation, and significantly reduced cycle times, but requires a proportional increase in sample load to overcome reduced ionisation efficiency at higher flow rates [183]. In MSi-TDP, the LC co-elution of proteoforms with multiple overlapping charge states reduces sensitivity as molecules of the same proteoform are spread across multiple charge states, yielding data that are challenging to interpret [60,184]. Capillary electrophoresis has the potential to completely resolve individual proteoforms from others, resulting in simplified MS spectra, but the incompatibility of solvents and buffers with ESI and the extremely small injection volumes required have restrained its routine implementation [185,186]. Additionally, CE cannot address the charge state loss in sensitivity and suffers from the same limitation as BUP and other MSi-TDP approaches in that any experimental replicates must be carried out sequentially as parallel replicates are not possible.

4.4. Improvements in Mass Spectrometry

Advances in instrumentation have all focused on the same outcome: increasing the dynamic range of concentration able to be analysed and accurately quantified through either an increase in scan speeds or further separation of LC co-eluting ions inside the MS. The most popular solution has been the application of Trapped Ion Mobility Spectrometry (TIMS), as applied in the Bruker TIMSToF platform where the TIMS device serves two purposes: (1) accumulation of ions of the same type to increase sensitivity; and (2) mobility-based separation of batches of these ions to enable another level of separation than that permitted by LC alone, thus reducing the diversity of m/z ions reaching the detector at any particular moment in time [187,188]. This has resulted in an increase in proteogenomic depth of coverage and some increase in individual ORF coverage [189]. Meanwhile, in the MSi-TDP space, single molecule detection could solve problems related to proteoforms taking on multiple charge states and confounding analysis [190], but it is not clear how quantification is achieved. What is clear is that vendors and researchers are still largely focused on peptide-centric analysis, which perpetuates BUP yet also enhances the power of iTDP to deeply analyse proteoforms and thus proteomes.

With developments in instrumentation comes the need for development of software and algorithms to identify the peptides and proteoforms contained within MS/MS spectra. Data analysis in BUP has traditionally had a 'flavour of the month' mentality tempered by a software package's ease of use and ease of integration into other analysis pipelines. MaxQuant [191] has long been the pipeline of choice because it is free, its output is readily accepted by pathway analysis pipelines [192], and it has a large community of users and resources available for those needing help. As computational resources have decreased in cost and developers have better understood how to make their software leverage those resources, BUP data analysis has moved to "Open Search" approaches in an attempt to assign more MS/MS spectra to a peptide sequence, especially those with PTM. Fragpipe [193] has been the most successful example of this, rapidly replacing MaxQuant as the pipeline of choice with downstream pathway pipelines being adapted to accept Fragpipe output [194].

The biggest change in BUP in the time of Proteomes' existence has been the adoption of Data-Independent Acquisition (DIA) [195], which attempts to overcome the stochastic selection of peptides for fragmentation used in Data-Dependent Acquisition (DDA) approaches, which can result in missing values. DIA should result in a data file that contains fragmentation of every peptide able to be ionised, enabling retrospective analysis of those data files. However, DIA has also led to numerous analysis conundrums that

were already a feature of BUP, including issues such as which peptide did a particular fragment ion belong to, thus simply deepening the protein inference problem. Ironically, successful DIA has been most reliant on comprehensive spectral libraries generated by DDA of highly fractionated peptide mixtures, although library-free methods such as that implemented in Progenesis [196] and DIA-NN [197] are in increasing use to overcome issues caused by Window-based DIA methods that do not measure intact peptide masses. Interestingly, library-free DIA has been a feature of Waters' mass spectrometers that are able to separate peptides by ion mobility prior to measurement of their intact m/z and of the fragments [198]. Fragments with the same mobility value as the parent peptide must have come from that peptide, allowing a spectrum to be generated without confounding fragments from other peptides. Unfortunately, data from Waters instruments are incompatible with pipelines such as Fragpipe and DIA-NN due to the unwillingness of their developers to fully meet the needs of users, thus funnelling researchers to certain instrument vendors. Regarding proteoforms, DIA is limited by whether peptides defining a proteoform (e.g., by carrying sequence variants or PTM) are present in a spectral library; however, this is not a common characteristic of libraries because of the aforementioned lack of detection of these peptides in BUP, even with extensive fractionation. iTDP (2DE/LC/TMS) analysis of proteoforms could provide the reference TMS spectra required for comprehensive spectral libraries that include proteoform-specific peptides for DIA.

4.5. Improvements in the Depth of Proteome Analysis

The release of Bruker's TIMSTOF platform also saw a resurgence in articles claiming to be performing 'deep' or 'comprehensive' proteome analysis [198–201]. As already emphasised, a lack of knowledge of the actual diversity of mature proteoforms in a cell makes claims of 'deep' proteome analysis rather pointless. 'Deep' proteome analysis is thus an especially troubling term when applied to single cell proteomics (SCP) [130,202–206]. The current 'State of the Art' reports ~3000–5000 proteins (ORF products) able to be identified and quantified [130]; this recent study does not provide the necessary detail for critical evaluation (i.e., how many canonical proteins identified with a single peptide; how many with 2 or 3 or more), noting only a median protein sequence coverage of 12.9% for single cells. It is, of course, the variance in data that informs on quality. Furthermore, the need to analyse hundreds of cells on a single instrument is resulting in relatively short LC/MS analysis times (30 min or less) [130], which will lead to compromises due to scan speed limitations and ion suppression of co-eluting peptides which would otherwise be further separated in time. In addition, all of the aforesaid issues with BUP still apply in SCP with reports not addressing the identification of proteoforms, although peptide speciation is beginning to be reported. For MSi-TDP, moving to single cells has necessitated the assessment of very large cells (i.e., muscle fibres) having only one or a few hyper-abundant species [207]. Again, this somehow seems to 'justify' the continued pursuit of BUP, but that simply leaves issues as they already exist and mires us in the promise of little advancement over 'standardised' proteogenomics over the next decade or more. There also appears to be little concern with sample preparation in regard to how the single cells are isolated and how that might affect the single cell proteome and thus how representative the findings are of the *in situ* native state. In this regard, while the issues associated with assessing cultured cell lines are more obvious, questions arise such as how does local heating during laser ablation [208] affect the proteomes of single cells close to and further from the line of excision? While it is not a question of how potentially important single cell analyses may be, much of the work has the quality of the technical attempts at *tour de force* studies [130] reminiscent of the first decade of BUP analyses and which still routinely appear in the literature [90]. But how much of that canonical cataloguing has effectively been turned into applicable knowledge? It is not what has been done that is important but rather what it means and thus what we learn from it. Perhaps the critical question should be 'how genuinely and quantitatively deep can these analyses be pushed with current and

developing technology’ rather than ‘how many canonical protein identifications can be quickly inferred with little if any knowledge of proteoforms’?

4.6. Developments in Alternative Proteome Analysis Technologies

There has always been a trend in science to adapt technologies from one field to another. Micro-arrays used for transcriptome analysis were adapted to protein arrays, although the technology did not achieve widespread use because of cost and other deficiencies [209,210]. Proteomics is currently seeing adoption of technology from genomics and NGS, where signal output is from amplification of oligonucleotide fragments attached to either antibodies (Olink) or aptamers (Somascan) that bind to a discrete part of a protein molecule, usually a series of amino acids (that could potentially include a PTM) [209,211–213]. While these technologies claim to be able to quantify up to 11,000 canonical proteins in a sample, they are subject to the same problem as protein arrays, that of being limited in their analysis to whatever “proteins” are targeted by the analysis panel. As we do not yet understand the diversity of proteoforms within an organism, it follows that these platforms also cannot measure specific proteoforms unless those are specifically targeted in the panel; thus, these analyses yield the same lack of proteoform detection and quantification as with BUP. This is not to say that these approaches have no value because there are reports of a >10 orders of magnitude dynamic range of detection, which is (minimally) estimated to be required for analysing samples such as serum or plasma. If the research question being asked genuinely fits within the limitations of these technologies, they may be effective tools to help uncover biology. However, as emphasised above, there is the need to subsequently pursue any finding to the proteoform level in order to identify the genuinely critical player(s).

The other emerging technology from genomics is nanopore sequencing, in which the passing of a linearised amino acid chain through a protein-based nanopore induces specific measurable electrical current changes depending on each particular amino acid [214,215]. There is recent evidence to suggest that PTM can also be identified and localised with this technology [216,217]. Issues to address include whether there are size limitations to the amino acid sequences that can be effectively linearised and ‘read’, and whether all (known) PTM can be distinguished or whether there will be overlapping and/or ambiguous signals for some.

Other technologies (e.g., refinements to Edman degradation, dendrimers, affinity matrices, BioID, DNA-PAINT, FRET X, CITE-seq and other RNA- and antibody-based techniques), in particular single molecule sequencing approaches, are in early stages but show promise, particularly when they demonstrably go beyond simply identifying canonical proteins but rather already address the need to analyse proteoforms [218]. Overall, there is potential in some of these approaches, but each has its own inherent technical limitations which are further complicated if proteoforms are the intended analytes for critical deep, *quantitative* analyses.

4.7. Integration with Other Omics Data

While proteoform-level analysis of the proteome is the necessary step forward that the field has to make, an equally important step is the integration of proteomics data with other omics data, especially metabolomics. Significant work has been done in this area with the creation of databases such as StringDB [219,220], PANTHER [221], and Reactome [222]; however, proteoform level data is lacking. Thus, although rarely done, researchers must consider and acknowledge the limitation that use of these software tools introduces in that this represents another form of inference since only canonical protein identifications are utilised. In their current state, obtaining usable data from these databases requires identified proteins to be submitted as gene names, as protein isoforms are not properly recognised. This is a curation problem that will only be solved by researchers submitting proteoform-defining information for review and inclusion and the database better recognising identifiers that define individual proteoforms (e.g., see suggested nomenclature [223]).

5. How to Move the Field More Rapidly Forward

Critically, broad recognition and exploitation of the complementarity of currently successful approaches is necessary, if not mandatory. Thus, for example, iTDP can quite effectively ‘fill the gaps’ until MSi-TDP has the necessary and sufficient capacity to fully address native proteomes across the full MW range of inherent proteoforms. Indeed, it has been known for over 20 years that TDP of larger species is best accomplished using an iTDP approach (at the time, in relation to MSi-TDP, inaptly named a ‘middle-down’ approach) [84]. Thus, by also utilising high-resolution 2DE as the front-end method to separate native proteoforms, current iTDP approaches most effectively capitalise on the complementary advances in both 2DE and BUP to enable genuinely deep, comprehensive proteome analyses. Rather than working largely in isolation from each other (if not actively against each other), which really has characterised much of proteomics over the last ~20+ years, the field must come to a transparent acceptance of the actual strengths and weakness of available methodologies and collaborate to capitalise on the former and in doing so, limit, if not eliminate, the latter. It is thus curious that MSi-TDP practitioners have avoided the obvious complementarity and improved analytical potential offered by 2DE. Furthermore, new technologies need to carefully continue through critical vetting processes that enable their ongoing evaluation as they are tested with increasingly complex samples and by several independent research teams in parallel. Such an approach will enable rapid addressing of issues rather than have them only become widely apparent years after full implementation of the method in the field. While that will certainly not provide an absolute guarantee that all problems will be identified in a timely fashion, it should significantly limit the now-usual pattern of having to address problems over the course of decades, which often leaves questionable data in the literature.

6. Consequences of a Failure to Address Proteoforms—the Price of Ignorance

The future lies in deeply understanding systems to identify specific and selective biomarkers and therapeutic targets. This is the only reasonable approach to genuine dissection of biological systems. Proteogenomics can not address the complexity of proteomes, although it might provide leads in those disorders having a direct genetic linkage; nonetheless, any potential leads must still then be pursued at the proteoform level. In this regard, for diseases resulting in abundance changes in a proteoform containing one or preferably more proteotypic peptides that can be subjected to targeted MS, this might provide a rapid diagnostic. But how often is this likely to be the case, noting that many of our most critical healthcare burdens are multi-factorial in nature? Furthermore, the current situation is that targeted MS studies (i.e., SRM/MRM; selected/multiple reaction monitoring) rarely target proteoforms beyond perhaps size variants (although exceptions are appearing [224]); appropriately using at least three peptides spanning the target species sequence is rarely done, and adding the complexity of specifically modified peptide standards to effectively calibrate the system for those PTM defining the proteoform of interest is a further demand and added expense. However, without these, specific proteoform identification and quantification is impossible.

Only a deep understanding of proteomes (and metabolomes, lipidomes, and transcriptomes) can provide the necessary functional and integrated understanding at the level of systems biology. The potential dangers of not deeply understanding the true functional components of systems—proteoforms—in an age in which techniques such as CRISPR are moving us ever closer to the realm of ‘routinely’ altering (defective) genes should be clear. How will a system that has developed without a specific functional protein/proteoform react to the expression of the ‘normal’ amino acid sequence? Will the system respond with (in)appropriate PTM? With any necessary PTM? The reality is that such treatments will (hopefully) target select cell types but these will reside within a whole system rather than the *in vitro* testing with specific cells in culture. We know already that, while generally effective, monoclonal antibody drugs are not entirely selective and thus even these therapeutics are not without side effects. How much more selective, and therefore perhaps

devoid of off-target effects, would drugs be, regardless of their nature/type, if they were targeted to a specific (offending) proteoform—and thus that specific resulting folded (i.e., 3D) species—rather than broadly against a canonical amino acid sequence? This identifies another link in the chain that needs to be better addressed. While there has been substantial progress in now being able to predict protein structures using machine learning ('AI') approaches, these are not without serious issues [225]. In large part, this is likely still due to the influences of proteogenomics and the notion that all that is needed is a canonical 'protein' identification. What the structural prediction approaches realistically need to focus on is proteoforms [226]; only such an approach will prove truly useful to, for example, drug development [227]. Coming full circle, this again requires that proteoforms become the standard target level of all proteomic analyses if the field is to move constructively forward and make real contributions to health, agriculture, and environmental issues.

It is time to escape the technique or technology-centric biases that have dominated the field for far too long. These somewhat ego-driven, blindered approaches serve only the status quo and/or the development of new business approaches that still focus on proteogenomics, albeit with evolved technologies [130]. We must accept that 'fitness-for-purpose' applies, and must utilise and/or integrate available methods accordingly [7]. Clearly, development must always continue, as must routine (re)evaluation of established methods with the aim of constant improvement [10].

Among the critical questions that arise with available approaches, perhaps the most important is this: what if different 'proteins' are identified as important (e.g., significantly changing in abundance between test conditions) by iTDP vs. BUP because the latter does not discriminate proteoforms? What is missed? What is over-emphasized in importance? Which approach is the more relevant if we take genuine proteome complexity into account [11]?

7. Being the Difference: The Proteomes Journal Approach

At Proteomes, our established publication policies/expectations preferentially take the longer view, that native complexity must ultimately be the focus, and thus routinely and aggressively addressed where methods enable such critical analyses. It is thus expected that the concept of proteoforms and/or proteome complexity be at least touched upon in every published paper, even if the methods used do not directly enable proteoform assessment. Authors are expected to transparently address the pros and cons of their study, again with the understanding that the complexity of proteomes must be acknowledged and how the work contributes or will contribute to furthering that understanding.

What the discipline of Proteomics no longer needs is self-appointed leaders but rather leadership and vision with a focus on genuinely addressing the real complexity of proteomes. We must recognise this as the post-proteogenomic era. The question thus arises as to how to future-proof the field from (ongoing) approaches that largely address only the low-hanging fruit of canonical amino acid sequences or only low MW species? Proteomes believes it is time to spearhead a more complete working definition of proteomes and encourage innovative approach(es) to effectively drive the field forward as critically and quantitatively as possible. It is time to look forward and fully embrace the genuine complexity of proteomes and what it will mean to routinely analyse them as deeply and quantitatively as possible.

8. Conclusions/Directions/Rationale

Some openly bemoan the fact that it is (increasingly) difficult to continuously secure funding for ever-newer mass spectrometers (i.e., the 'keeping-up-with-the-Joneses' problem). Perhaps what funding agencies are/should be looking for is an effective analytical approach that provides the rigorous biological information necessary to understand and effectively target/dissect molecular mechanisms, and thereby identify rational new drug targets as well as biomarkers. Such rigorously identified therapeutics and biomarkers can and will subsequently survive appropriately rigorous validation, including clinical trials.

Considering that ~86% of drug candidates between 2000–2015 failed in clinical trials, representing an exorbitant cost in both time and money [228], it is time to accept that ‘traditional’ approaches no longer suffice. Only the routine, deep proteoform level assessments provided by iTDP will yield critical systems biology knowledge. In this regard, it is surprising that LC/mass spectrometer manufacturers have not sought to offer more rigorous front-end analytical tools (i.e., 2DE and high-end imaging instrumentation) to best complement and capitalise on the LC/TMS equipment already marketed, and thus most effectively address the full needs of proteomics research [10]. Nonetheless, the future is promising, considering that rigorous iTDP approaches are well supported even by older (and less expensive) MS systems. That said, ion mobility MS may well prove to be a powerful tool in addressing proteoforms [229]. The ongoing development of nanopore approaches also appears promising in terms of potentially being capable of quantitatively assessing the full complement of proteoforms in a biological extract. While acknowledging this potential, enough questions remain, most specifically concerning proteoform complexity, that it seems unlikely that this approach will see wide-scale application to whole proteomes in the near immediate future, although one might imagine that quantitative targeted applications could appear at a reasonable pace. Direct elution of resolved proteoforms from 2D gel spots into a nanopore device might prove particularly advantageous. A critical focus on interactions—noting, however, that current widely used software applications, such as STRINGDB and PANTHER, address only canonical proteins and do not discern specific proteoform functions from the literature (if such information is even available)—will be essential to our understanding of systems at a genuinely functional level. This, then, also emphasises the need for better (1) structural analyses and predictions, that focus on proteoforms rather than only canonical amino acid sequences; (2) spatial resolution (e.g., in MS imaging); (3) temporal resolution (e.g., to assess transient proteoforms, perhaps even in signalling networks) [230]; (4) understanding the implications of PTM crosstalk [231]; and careful consideration of (4) the potential applications of machine learning to addressing data analysis and interpretation [232]. With such routine deep proteome analyses comes the very real promise of far more selective biomarkers, drug targets, and personalised—or even realistically individualised—medicine. Only a rigorous focus on analytical quality will get us there.

Author Contributions: All authors contributed equally and have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. O’Farrell, P.H. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **1975**, *250*, 4007–4021. [[CrossRef](#)] [[PubMed](#)]
2. Pieper, R.; Gatlin, C.L.; Makusky, A.J.; Russo, P.S.; Schatz, C.R.; Miller, S.S.; Su, Q.; McGrath, A.M.; Estock, M.A.; Parmar, P.P.; et al. The human serum proteome: Display of nearly 3700 chromatographically separated protein spots on two-dimensional electrophoresis gels and identification of 325 distinct proteins. *Proteomics* **2003**, *3*, 1345–1364. [[CrossRef](#)] [[PubMed](#)]
3. Thiede, B.; Koehler, C.J.; Strozynski, M.; Treumann, A.; Stein, R.; Zimny-Arndt, U.; Schmid, M.; Jungblut, P.R. High resolution quantitative proteomics of HeLa cells protein species using stable isotope labeling with amino acids in cell culture (SILAC), two-dimensional gel electrophoresis (2DE) and nano-liquid chromatography coupled to an LTQ-Orbitrap Mass spectrometer. *Mol. Cell Proteom.* **2013**, *12*, 529–538. [[CrossRef](#)]
4. Wright, E.P.; Partridge, M.A.; Padula, M.P.; Gauci, V.J.; Malladi, C.S.; Coorsen, J.R. Top-down proteomics: Enhancing 2D gel electrophoresis from tissue processing to high-sensitivity protein detection. *Proteomics* **2014**, *14*, 872–889. [[CrossRef](#)] [[PubMed](#)]
5. Naryzhny, S. Towards the Full Realization of 2DE Power. *Proteomes* **2016**, *4*, 33. [[CrossRef](#)] [[PubMed](#)]

6. Zhan, X.; Yang, H.; Peng, F.; Li, J.; Mu, Y.; Long, Y.; Cheng, T.; Huang, Y.; Li, Z.; Lu, M.; et al. How many proteins can be identified in a 2DE gel spot within an analysis of a complex human cancer tissue proteome? *Electrophoresis* **2018**, *39*, 965–980. [[CrossRef](#)] [[PubMed](#)]
7. Coorssen, J.R.; Yergey, A.L. Proteomics Is Analytical Chemistry: Fitness-for-Purpose in the Application of Top-Down and Bottom-Up Analyses. *Proteomes* **2015**, *3*, 440–453. [[CrossRef](#)]
8. Oliveira, B.M.; Coorssen, J.R.; Martins-de-Souza, D. 2DE: The phoenix of proteomics. *J. Proteom.* **2014**, *104*, 140–150. [[CrossRef](#)] [[PubMed](#)]
9. Coorssen, J.; Yergey, A. Editorial for Special Issue: Approaches to Top-Down Proteomics: In Honour of Prof. Patrick H. O’Farrell. *Proteomes* **2017**, *5*, 18. [[CrossRef](#)]
10. Carbonara, K.; Andonovski, M.; Coorssen, J.R. Proteomes Are of Proteoforms: Embracing the Complexity. *Proteomes* **2021**, *9*, 38. [[CrossRef](#)]
11. Ercan, H.; Resch, U.; Hsu, F.; Mitulovic, G.; Bileck, A.; Gerner, C.; Yang, J.-W.; Geiger, M.; Miller, I.; Zellner, M. A Practical and Analytical Comparative Study of Gel-Based Top-Down and Gel-Free Bottom-Up Proteomics Including Unbiased Proteoform Detection. *Cells* **2023**, *12*, 747. [[CrossRef](#)] [[PubMed](#)]
12. Mørtz, E.; Vorm, O.; Mann, M.; Roepstorff, P. Identification of proteins in polyacrylamide gels by mass spectrometric peptide mapping combined with database search. *Biol. Mass Spectrom.* **1994**, *23*, 249–261. [[CrossRef](#)] [[PubMed](#)]
13. Matsumoto, H.; Kurien, B.T.; Takagi, Y.; Kahn, E.S.; Kinumi, T.; Komori, N.; Yamada, T.; Hayashi, F.; Isono, K.; Pak, W.L.; et al. Phosrestin I undergoes the earliest light-induced phosphorylation by a calcium/calmodulin-dependent protein kinase in drosophila photoreceptors. *Neuron* **1994**, *12*, 997–1010. [[CrossRef](#)] [[PubMed](#)]
14. Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Anal. Chem.* **1996**, *68*, 850–858. [[CrossRef](#)] [[PubMed](#)]
15. Wilm, M.; Shevchenko, A.; Houthaeve, T.; Breit, S.; Schweigerer, L.; Fotsis, T.; Mann, M. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* **1996**, *379*, 466–469. [[CrossRef](#)]
16. Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. *Nature* **2003**, *422*, 198–207. [[CrossRef](#)]
17. Coorssen, J. Analytical Approaches to Address Proteome Complexity. *Biotechniques* **2023**. Available online: <https://www.biotechniques.com/proteomics/ebook-lab-essentials-proteomics/> (accessed on 16 April 2024).
18. Eng, J.K.; McCormack, A.L.; Yates, J.R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976–989. [[CrossRef](#)]
19. Link, A.J.; Eng, J.; Schieltz, D.M.; Carmack, E.; Mize, G.J.; Morris, D.R.; Garvik, B.M.; Yates, J.R. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **1999**, *17*, 676–682. [[CrossRef](#)]
20. Washburn, M.P.; Wolters, D.; Yates, J.R., 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001**, *19*, 242–247. [[CrossRef](#)]
21. Beasley-Green, A.; Heckert, N.A. Estimation of measurement uncertainty for the quantification of protein by ID-LC-MS/MS. *Anal. Bioanal. Chem.* **2023**, *415*, 3265–3274. [[CrossRef](#)]
22. Dupree, E.J.; Jayathirtha, M.; Yorkey, H.; Mihasan, M.; Petre, B.A.; Darie, C.C. A Critical Review of Bottom-Up Proteomics: The Good, the Bad, and the Future of This Field. *Proteomes* **2020**, *8*, 14. [[CrossRef](#)] [[PubMed](#)]
23. Orsburn, B.C. Evaluation of the Sensitivity of Proteomics Methods Using the Absolute Copy Number of Proteins in a Single Cell as a Metric. *Proteomes* **2021**, *9*, 34. [[CrossRef](#)] [[PubMed](#)]
24. Aziz, S.; Rasheed, F.; Zahra, R.; König, S. Mass Spectrometry-Based Proteomics of Minor Species in the Bulk: Questions to Raise with Respect to the Untargeted Analysis of Viral Proteins in Human Tissue. *Life* **2023**, *13*, 544. [[CrossRef](#)] [[PubMed](#)]
25. Schork, K.; Podwojski, K.; Turewicz, M.; Stephan, C.; Eisenacher, M. Important Issues in Planning a Proteomics Experiment Planning a Proteomics experiment: Statistical Considerations of Quantitative Proteomic Quantitative proteomics Data. In *Quantitative Methods in Proteomics*; Marcus, K., Eisenacher, M., Sitek, B., Eds.; Springer US: New York, NY, USA, 2021; pp. 1–20.
26. Prakash, A.; Rezai, T.; Krastins, B.; Sarracino, D.; Athanas, M.; Russo, P.; Zhang, H.; Tian, Y.; Li, Y.; Kulasingam, V.; et al. Interlaboratory reproducibility of selective reaction monitoring assays using multiple upfront analyte enrichment strategies. *J. Proteome Res.* **2012**, *11*, 3986–3995. [[CrossRef](#)] [[PubMed](#)]
27. Collins, B.C.; Hunter, C.L.; Liu, Y.; Schilling, B.; Rosenberger, G.; Bader, S.L.; Chan, D.W.; Gibson, B.W.; Gingras, A.C.; Held, J.M.; et al. Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nat. Commun.* **2017**, *8*, 291. [[CrossRef](#)] [[PubMed](#)]
28. Tabb, D.L.; Vega-Montoto, L.; Rudnick, P.A.; Variyath, A.M.; Ham, A.J.; Bunk, D.M.; Kilpatrick, L.E.; Billheimer, D.D.; Blackman, R.K.; Cardasis, H.L.; et al. Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2010**, *9*, 761–776. [[CrossRef](#)] [[PubMed](#)]
29. Poulos, R.C.; Hains, P.G.; Shah, R.; Lucas, N.; Xavier, D.; Manda, S.S.; Anees, A.; Koh, J.M.S.; Mahboob, S.; Wittman, M.; et al. Strategies to enable large-scale proteomics for reproducible research. *Nat. Commun.* **2020**, *11*, 3793. [[CrossRef](#)] [[PubMed](#)]
30. Uszkoreit, J.; Palmblad, M.; Schwämmle, V. Tackling reproducibility: Lessons for the proteomics community. *Expert Rev. Proteom.* **2024**, *21*, 9–11. [[CrossRef](#)]
31. Bludau, I.; Frank, M.; Dorig, C.; Cai, Y.; Heusel, M.; Rosenberger, G.; Picotti, P.; Collins, B.C.; Rost, H.; Aebersold, R. Systematic detection of functional proteoform groups from bottom-up proteomic datasets. *Nat. Commun.* **2021**, *12*, 3810. [[CrossRef](#)]

32. Plubell, D.L.; Kall, L.; Webb-Robertson, B.J.; Bramer, L.M.; Ives, A.; Kelleher, N.L.; Smith, L.M.; Montine, T.J.; Wu, C.C.; MacCoss, M.J. Putting Humpty Dumpty Back Together Again: What Does Protein Quantification Mean in Bottom-Up Proteomics? *J. Proteome Res.* **2022**, *21*, 891–898. [[CrossRef](#)]
33. Jungblut, P.R.; Holzhtter, H.G.; Apweiler, R.; Schluter, H. The speciation of the proteome. *Chem. Cent. J.* **2008**, *2*, 16. [[CrossRef](#)] [[PubMed](#)]
34. Jungblut, P.R.; Thiede, B.; Schluter, H. Towards deciphering proteomes via the proteoform, protein speciation, moonlighting and protein code concepts. *J. Proteom.* **2016**, *134*, 1–4. [[CrossRef](#)] [[PubMed](#)]
35. Vanderperre, B.; Lucier, J.-F.; Bissonnette, C.; Motard, J.; Tremblay, G.; Vanderperre, S.; Wisztorski, M.; Salzet, M.; Boisvert, F.-M.; Roucou, X. Direct Detection of Alternative Open Reading Frames Translation Products in Human Significantly Expands the Proteome. *PLoS ONE* **2013**, *8*, e70698. [[CrossRef](#)] [[PubMed](#)]
36. Coorssen, J. Why a ‘Protein’ Isn’t: Acknowledging Proteome Complexity. *Biotechniques* **2023**. Available online: <https://www.biotechniques.com/proteomics/why-a-protein-isnt-acknowledging-proteome-complexity/> (accessed on 16 April 2024).
37. Mellacheruvu, D.; Wright, Z.; Couzens, A.L.; Lambert, J.-P.; St-Denis, N.A.; Li, T.; Miteva, Y.V.; Hauri, S.; Sardi, M.E.; Low, T.Y.; et al. The CRAPome: A contaminant repository for affinity purification–mass spectrometry data. *Nat. Methods* **2013**, *10*, 730–736. [[CrossRef](#)] [[PubMed](#)]
38. Su, J.; Yang, L.; Sun, Z.; Zhan, X. Personalized Drug Therapy: Innovative Concept Guided with Proteoforms. *Mol. Cell Proteom.* **2024**, *23*, 100737. [[CrossRef](#)]
39. Wilczek, F. Einstein’s Parable of Quantum Insanity. *Quanta Magazine*. 2015. Available online: <https://www.scientificamerican.com/article/einstein-s-parable-of-quantum-insanity/> (accessed on 16 April 2024).
40. Aebersold, R.; Agar, J.N.; Amster, I.J.; Baker, M.S.; Bertozzi, C.R.; Boja, E.S.; Costello, C.E.; Cravatt, B.F.; Fenselau, C.; Garcia, B.A.; et al. How many human proteoforms are there? *Nat. Chem. Biol.* **2018**, *14*, 206–214. [[CrossRef](#)] [[PubMed](#)]
41. Chen, J.; Brunner, A.-D.; Cogan, J.Z.; Nuñez, J.K.; Fields, A.P.; Adamson, B.; Itzhak, D.N.; Li, J.Y.; Mann, M.; Leonetti, M.D.; et al. Pervasive functional translation of noncanonical human open reading frames. *Science* **2020**, *367*, 1140–1146. [[CrossRef](#)] [[PubMed](#)]
42. Cao, X.; Khitun, A.; Na, Z.; Dumitrescu, D.G.; Kubica, M.; Olatunji, E.; Slavoff, S.A. Comparative Proteomic Profiling of Unannotated Microproteins and Alternative Proteins in Human Cell Lines. *J. Proteome Res.* **2020**, *19*, 3418–3426. [[CrossRef](#)]
43. Schlesinger, D.; Elsässer, S.J. Revisiting sORFs: Overcoming challenges to identify and characterize functional microproteins. *FEBS J.* **2022**, *289*, 53–74. [[CrossRef](#)]
44. Harper, J.W.; Bennett, E.J. Proteome complexity and the forces that drive proteome imbalance. *Nature* **2016**, *537*, 328–338. [[CrossRef](#)]
45. Rappsilber, J.; Mann, M. What does it mean to identify a protein in proteomics? *Trends Biochem. Sci.* **2002**, *27*, 74–78. [[CrossRef](#)]
46. Rocha, J.J.; Jayaram, S.A.; Stevens, T.J.; Muschalik, N.; Shah, R.D.; Emran, S.; Robles, C.; Freeman, M.; Munro, S. Functional unknowns: Systematic screening of conserved genes of unknown function. *PLoS Biol.* **2023**, *21*, e3002222. [[CrossRef](#)] [[PubMed](#)]
47. Kustatscher, G.; Collins, T.; Gingras, A.-C.; Guo, T.; Hermjakob, H.; Ideker, T.; Lilley, K.S.; Lundberg, E.; Marcotte, E.M.; Ralser, M.; et al. Understudied proteins: Opportunities and challenges for functional proteomics. *Nat. Methods* **2022**, *19*, 774–779. [[CrossRef](#)]
48. Faria, S.S.; Morris, C.F.M.; Silva, A.R.; Fonseca, M.P.; Forget, P.; Castro, M.S.; Fontes, W. A Timely Shift from Shotgun to Targeted Proteomics and How It Can Be Groundbreaking for Cancer Research. *Front. Oncol.* **2017**, *7*, 13. [[CrossRef](#)]
49. Henzel, W.J.; Billeci, T.M.; Stults, J.T.; Wong, S.C.; Grimley, C.; Watanabe, C. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 5011–5015. [[CrossRef](#)]
50. Büttner, K.; Bernhardt, J.; Scharf, C.; Schmid, R.; Mäder, U.; Eymann, C.; Antelmann, H.; Völker, A.; Völker, U.; Hecker, M. A comprehensive two-dimensional map of cytosolic proteins of *Bacillus subtilis*. *Electrophoresis* **2001**, *22*, 2908–2935. [[CrossRef](#)]
51. Tonella, L.; Hoogland, C.; Binz, P.A.; Appel, R.D.; Hochstrasser, D.F.; Sanchez, J.C. New perspectives in the *Escherichia coli* proteome investigation. *Proteomics* **2001**, *1*, 409–423. [[CrossRef](#)]
52. Marcus, K.; Immler, D.; Sternberger, J.; Meyer, H.E. Identification of platelet proteins separated by two-dimensional gel electrophoresis and analyzed by matrix assisted laser desorption/ionization-time of flight-mass spectrometry and detection of tyrosine-phosphorylated proteins. *Electrophoresis* **2000**, *21*, 2622–2636. [[CrossRef](#)] [[PubMed](#)]
53. Raymackers, J.; Daniels, A.; De Brabandere, V.; Missiaen, C.; Dauwe, M.; Verhaert, P.; Vanmechelen, E.; Meheus, L. Identification of two-dimensionally separated human cerebrospinal fluid proteins by N-terminal sequencing, matrix-assisted laser desorption/ionization—mass spectrometry, nanoliquid chromatography-electrospray ionization-time of flight-mass spectrometry, and tandem mass spectrometry. *Electrophoresis* **2000**, *21*, 2266–2283. [[CrossRef](#)] [[PubMed](#)]
54. Hunsucker, S.W.; Duncan, M.W. Is protein overlap in two-dimensional gels a serious practical problem? *Proteomics* **2006**, *6*, 1374–1375. [[CrossRef](#)]
55. Zhan, X.; Li, B.; Zhan, X.; Schlüter, H.; Jungblut, P.R.; Coorssen, J.R. Innovating the Concept and Practice of Two-Dimensional Gel Electrophoresis in the Analysis of Proteomes at the Proteoform Level. *Proteomes* **2019**, *7*, 36. [[CrossRef](#)] [[PubMed](#)]
56. Wasinger, V.C.; Cordwell, S.J.; Cerpa-Poljak, A.; Yan, J.X.; Gooley, A.A.; Wilkins, M.R.; Duncan, M.W.; Harris, R.; Williams, K.L.; Humphery-Smith, I. Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis* **1995**, *16*, 1090–1094. [[CrossRef](#)] [[PubMed](#)]
57. Cody, R.B.; Amster, I.J.; McLafferty, F.W. Peptide mixture sequencing by tandem Fourier-transform mass spectrometry. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 6367–6370. [[CrossRef](#)] [[PubMed](#)]

58. Kelleher, N.L.; Lin, H.Y.; Valaskovic, G.A.; Aaserud, D.J.; Fridriksson, E.K.; McLafferty, F.W. Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *J. Am. Chem. Soc.* **1999**, *121*, 806–812. [[CrossRef](#)]
59. Durbin, K.R.; Fornelli, L.; Fellers, R.T.; Doubleday, P.F.; Narita, M.; Kelleher, N.L. Quantitation and Identification of Thousands of Human Proteoforms below 30 kDa. *J. Proteome Res.* **2016**, *15*, 976–982. [[CrossRef](#)]
60. Huguet, R.; Mullen, C.; Srzentić, K.; Greer, J.B.; Fellers, R.T.; Zabrouskov, V.; Syka, J.E.P.; Kelleher, N.L.; Fornelli, L. Proton Transfer Charge Reduction Enables High-Throughput Top-Down Analysis of Large Proteoforms. *Anal. Chem.* **2019**, *91*, 15732–15739. [[CrossRef](#)]
61. Po, A.; Evers, C.E. Top-Down Proteomics and the Challenges of True Proteoform Characterization. *J. Proteome Res.* **2023**, *22*, 3663–3675. [[CrossRef](#)] [[PubMed](#)]
62. Tabb, D.L.; Jeong, K.; Druart, K.; Gant, M.S.; Brown, K.A.; Nicora, C.; Zhou, M.; Couvillion, S.; Nakayasu, E.; Williams, J.E.; et al. Comparing Top-Down Proteoform Identification: Deconvolution, PrSM Overlap, and PTM Detection. *J. Proteome Res.* **2023**, *22*, 2199–2217. [[CrossRef](#)]
63. Chen, B.; Brown, K.A.; Lin, Z.; Ge, Y. Top-Down Proteomics: Ready for Prime Time? *Anal. Chem.* **2018**, *90*, 110–127. [[CrossRef](#)]
64. Cui, W.; Rohrs, H.W.; Gross, M.L. Top-down mass spectrometry: Recent developments, applications and perspectives. *Analyst* **2011**, *136*, 3854–3864. [[CrossRef](#)]
65. Biemann, K.; Gapp, G.; Seibl, J. Application of Mass Spectrometry to Structure Problems. I. Amino Acid Sequence in Peptides. *J. Am. Chem. Soc.* **1959**, *81*, 2274–2275. [[CrossRef](#)]
66. Senn, M.; McLafferty, F.W. Automatic amino-acid-sequence determination in peptides. *Biochem. Biophys. Res. Commun.* **1966**, *23*, 381–385. [[CrossRef](#)] [[PubMed](#)]
67. Lucas, F.; Barber, M.; Wolstenholme, W.A.; Geddes, A.J.; Graham, G.N.; Morris, H.R. Mass-spectrometric determination of the amino acid sequences in peptides isolated from the protein silk fibroin of *Bombyx mori*. *Biochem. J.* **1969**, *114*, 695–702. [[CrossRef](#)] [[PubMed](#)]
68. Morris, H.R.; Geddes, A.J.; Graham, G.N. Some problems associated with the amino acid-sequence analysis of proteins by mass spectrometry. *Biochem. J.* **1969**, *111*, 38. [[CrossRef](#)] [[PubMed](#)]
69. Barber, M.; Powers, P.; Wallington, M.J.; Wolstenholme, W.A. Computer Interpretation of High Resolution Mass Spectra*. *Nature* **1966**, *212*, 784–787. [[CrossRef](#)] [[PubMed](#)]
70. Dayhoff, M.O.; Eck, R.V. MASSPEC: A computer program for complete sequence analysis of large proteins from mass spectrometry data of a single sample. *Comput. Biol. Med.* **1970**, *1*, 5–28. [[CrossRef](#)] [[PubMed](#)]
71. Loo, J.A.; Edmonds, C.G.; Smith, R.D. Primary sequence information from intact proteins by electrospray ionization tandem mass spectrometry. *Science* **1990**, *248*, 201–204. [[CrossRef](#)] [[PubMed](#)]
72. Loo, J.A.; Edmonds, C.G.; Smith, R.D. Tandem mass spectrometry of very large molecules: Serum albumin sequence information from multiply charged ions formed by electrospray ionization. *Anal. Chem.* **1991**, *63*, 2488–2499. [[CrossRef](#)]
73. Little, D.P.; Speir, J.P.; Senko, M.W.; O'Connor, P.B.; McLafferty, F.W. Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. *Anal. Chem.* **1994**, *66*, 2809–2815. [[CrossRef](#)]
74. Han, X.; Jin, M.; Breuker, K.; McLafferty, F.W. Extending Top-Down Mass Spectrometry to Proteins with Masses Greater Than 200 Kilodaltons. *Science* **2006**, *314*, 109–112. [[CrossRef](#)]
75. Michel, H.; Hunt, D.F.; Shabanowitz, J.; Bennett, J. Tandem mass spectrometry reveals that three photosystem II proteins of spinach chloroplasts contain N-acetyl-O-phosphothreonine at their NH₂ termini. *J. Biol. Chem.* **1988**, *263*, 1123–1130. [[CrossRef](#)] [[PubMed](#)]
76. Mørtz, E.; O'Connor, P.B.; Roepstorff, P.; Kelleher, N.L.; Wood, T.D.; McLafferty, F.W.; Mann, M. Sequence tag identification of intact proteins by matching tandem mass spectral data against sequence data bases. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8264–8267. [[CrossRef](#)]
77. Li, W.; Hendrickson, C.L.; Emmett, M.R.; Marshall, A.G. Identification of Intact Proteins in Mixtures by Alternated Capillary Liquid Chromatography Electrospray Ionization and LC ESI Infrared Multiphoton Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **1999**, *71*, 4397–4402. [[CrossRef](#)] [[PubMed](#)]
78. Skinner, O.S.; Haverland, N.A.; Fornelli, L.; Melani, R.D.; Do Vale, L.H.F.; Seckler, H.S.; Doubleday, P.F.; Schachner, L.F.; Srzentić, K.; Kelleher, N.L.; et al. Top-down characterization of endogenous protein complexes with native proteomics. *Nat. Chem. Biol.* **2018**, *14*, 36–41. [[CrossRef](#)] [[PubMed](#)]
79. Durbin, K.R.; Robey, M.T.; Voong, L.N.; Fellers, R.T.; Lutomski, C.A.; El-Baba, T.J.; Robinson, C.V.; Kelleher, N.L. ProSight Native: Defining Protein Complex Composition from Native Top-Down Mass Spectrometry Data. *J. Proteome Res.* **2023**, *22*, 2660–2668. [[CrossRef](#)] [[PubMed](#)]
80. Lloyd-Jones, C.; dos Santos Seckler, H.; DiStefano, N.; Sniderman, A.; Compton, P.D.; Kelleher, N.L.; Wilkins, J.T. Preparative Electrophoresis for HDL Particle Size Separation and Intact-Mass Apolipoprotein Proteoform Analysis. *J. Proteome Res.* **2023**, *22*, 1455–1465. [[CrossRef](#)] [[PubMed](#)]
81. Fornelli, L.; Durbin, K.R.; Fellers, R.T.; Early, B.P.; Greer, J.B.; LeDuc, R.D.; Compton, P.D.; Kelleher, N.L. Advancing Top-down Analysis of the Human Proteome Using a Benchtop Quadrupole-Orbitrap Mass Spectrometer. *J. Proteome Res.* **2017**, *16*, 609–618. [[CrossRef](#)] [[PubMed](#)]
82. Tran, J.C.; Doucette, A.A. Gel-Eluted Liquid Fraction Entrapment Electrophoresis: An Electrophoretic Method for Broad Molecular Weight Range Proteome Separation. *Anal. Chem.* **2008**, *80*, 1568–1573. [[CrossRef](#)]

83. Lee, J.E.; Kellie, J.F.; Tran, J.C.; Tipton, J.D.; Catherman, A.D.; Thomas, H.M.; Ahlf, D.R.; Durbin, K.R.; Vellaichamy, A.; Ntai, I.; et al. A robust two-dimensional separation for top-down tandem mass spectrometry of the low-mass proteome. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2183–2191. [[CrossRef](#)]
84. Forbes, A.J.; Mazur, M.T.; Kelleher, N.L.; Patel, H.M.; Walsh, C.T. Toward Efficient Analysis of >70 kDa Proteins with 100% Sequence Coverage. *Eur. J. Mass Spectrom.* **2001**, *7*, 81–87. [[CrossRef](#)]
85. Okkels, L.M.; Müller, E.-C.; Schmid, M.; Rosenkrands, I.; Kaufmann, S.H.E.; Andersen, P.; Jungblut, P.R. CFP10 discriminates between nonacetylated and acetylated ESAT-6 of Mycobacterium tuberculosis by differential interaction. *Proteomics* **2004**, *4*, 2954–2960. [[CrossRef](#)] [[PubMed](#)]
86. Meyer, B.; Papisotiriou, D.G.; Karas, M. 100% protein sequence coverage: A modern form of surrealism in proteomics. *Amino Acids* **2011**, *41*, 291–310. [[CrossRef](#)] [[PubMed](#)]
87. Frese, C.K.; Altalear, A.F.M.; van den Toorn, H.; Nolting, D.; Griep-Raming, J.; Heck, A.J.R.; Mohammed, S. Toward Full Peptide Sequence Coverage by Dual Fragmentation Combining Electron-Transfer and Higher-Energy Collision Dissociation Tandem Mass Spectrometry. *Anal. Chem.* **2012**, *84*, 9668–9673. [[CrossRef](#)] [[PubMed](#)]
88. Sinitcyn, P.; Richards, A.L.; Weatheritt, R.J.; Brademan, D.R.; Marx, H.; Shishkova, E.; Meyer, J.G.; Hebert, A.S.; Westphall, M.S.; Blencowe, B.J.; et al. Global detection of human variants and isoforms by deep proteome sequencing. *Nat. Biotechnol.* **2023**, *41*, 1776–1786. [[CrossRef](#)] [[PubMed](#)]
89. Schweiger-Hufnagel, U.; Hufnagel, P.; Hebel, R.; Witt, M.; Schmit, P.O.; Macht, M.; Asperger, A. Towards 100% Sequence Coverage in Protein QC: In-depth Sequence Characterization of Monoclonal Antibodies. *J. Biomol. Technol.* **2010**, *21*, S63.
90. Messner, C.B.; Demichev, V.; Bloomfield, N.; Yu, J.S.L.; White, M.; Kreidl, M.; Egger, A.-S.; Freiwald, A.; Ivosev, G.; Wasim, F.; et al. Ultra-fast proteomics with Scanning SWATH. *Nat. Biotechnol.* **2021**, *39*, 846–854. [[CrossRef](#)] [[PubMed](#)]
91. Ahrné, E.; Molzahn, L.; Glatter, T.; Schmidt, A. Critical assessment of proteome-wide label-free absolute abundance estimation strategies. *Proteomics* **2013**, *13*, 2567–2578. [[CrossRef](#)]
92. Dowell, J.A.; Wright, L.J.; Armstrong, E.A.; Denu, J.M. Benchmarking Quantitative Performance in Label-Free Proteomics. *ACS Omega* **2021**, *6*, 2494–2504. [[CrossRef](#)] [[PubMed](#)]
93. Sánchez, B.J.; Lahtvee, P.J.; Campbell, K.; Kasvandik, S.; Yu, R.; Domenzain, I.; Zelezniak, A.; Nielsen, J. Benchmarking accuracy and precision of intensity-based absolute quantification of protein abundances in *Saccharomyces cerevisiae*. *Proteomics* **2021**, *21*, e2000093. [[CrossRef](#)]
94. Millán-Oropeza, A.; Blein-Nicolas, M.; Monnet, V.; Zivy, M.; Henry, C. Comparison of Different Label-Free Techniques for the Semi-Absolute Quantification of Protein Abundance. *Proteomes* **2022**, *10*, 2. [[CrossRef](#)]
95. Huang, T.; Wang, J.; Yu, W.; He, Z. Protein inference: A review. *Brief. Bioinform.* **2012**, *13*, 586–614. [[CrossRef](#)] [[PubMed](#)]
96. Hamid, Z.; Zimmerman, K.D.; Guillen-Ahlers, H.; Li, C.; Nathanielsz, P.; Cox, L.A.; Olivier, M. Assessment of label-free quantification and missing value imputation for proteomics in non-human primates. *BMC Genom.* **2022**, *23*, 496. [[CrossRef](#)] [[PubMed](#)]
97. Kong, W.; Hui, H.W.H.; Peng, H.; Goh, W.W.B. Dealing with missing values in proteomics data. *Proteomics* **2022**, *22*, e2200092. [[CrossRef](#)] [[PubMed](#)]
98. Reddy, P.J.; Ray, S.; Srivastava, S. The quest of the human proteome and the missing proteins: Digging deeper. *OMICS* **2015**, *19*, 276–282. [[CrossRef](#)] [[PubMed](#)]
99. O'Brien, J.J.; Gunawardena, H.P.; Paulo, J.A.; Chen, X.; Ibrahim, J.G.; Gygi, S.P.; Qaqish, B.F. The effects of nonignorable missing data on label-free mass spectrometry proteomics experiments. *Ann. Appl. Stat.* **2018**, *12*, 2075–2095. [[CrossRef](#)] [[PubMed](#)]
100. Chen, X.; Sun, Y.; Zhang, T.; Shu, L.; Roepstorff, P.; Yang, F. Quantitative Proteomics Using Isobaric Labeling: A Practical Guide. *Genom. Proteom. Bioinform.* **2021**, *19*, 689–706. [[CrossRef](#)] [[PubMed](#)]
101. Hutchinson-Bunch, C.; Sanford, J.A.; Hansen, J.R.; Gritsenko, M.A.; Rodland, K.D.; Piehowski, P.D.; Qian, W.-J.; Adkins, J.N. Assessment of TMT Labeling Efficiency in Large-Scale Quantitative Proteomics: The Critical Effect of Sample pH. *ACS Omega* **2021**, *6*, 12660–12666. [[CrossRef](#)] [[PubMed](#)]
102. Padula, M.P.; Berry, I.J.; O'Rourke, M.B.; Raymond, B.B.; Santos, J.; Djordjevic, S.P. A Comprehensive Guide for Performing Sample Preparation and Top-Down Protein Analysis. *Proteomes* **2017**, *5*, 11. [[CrossRef](#)] [[PubMed](#)]
103. O'Rourke, M.B.; Town, S.E.L.; Dalla, P.V.; Bicknell, F.; Koh Belic, N.; Violi, J.P.; Steele, J.R.; Padula, M.P. What is Normalization? The Strategies Employed in Top-Down and Bottom-Up Proteome Analysis Workflows. *Proteomes* **2019**, *7*, 29. [[CrossRef](#)]
104. Meleady, P. *Proteomics Mass Spectrometry Methods: Sample Preparation, Protein Digestion, and Research Protocols*; Elsevier Science & Technology: San Diego, CA, USA, 2024.
105. Ignjatovic, V.; Geyer, P.E.; Palaniappan, K.K.; Chaaban, J.E.; Omenn, G.S.; Baker, M.S.; Deutsch, E.W.; Schwenk, J.M. Mass Spectrometry-Based Plasma Proteomics: Considerations from Sample Collection to Achieving Translational Data. *J. Proteome Res.* **2019**, *18*, 4085–4097. [[CrossRef](#)]
106. Huang, J.; Khademi, M.; Lindhe, Ö.; Jönsson, G.; Piehl, F.; Olsson, T.; Kockum, I. Assessing the Preanalytical Variability of Plasma and Cerebrospinal Fluid Processing and Its Effects on Inflammation-Related Protein Biomarkers. *Mol. Cell. Proteom.* **2021**, *20*, 100157. [[CrossRef](#)] [[PubMed](#)]
107. Plebani, M.; Banfi, G.; Bernardini, S.; Bondanini, F.; Conti, L.; Dorizzi, R.; Ferrara, F.E.; Mancini, R.; Trenti, T. Serum or plasma? An old question looking for new answers. *Clin. Chem. Lab. Med.* **2020**, *58*, 178–187. [[CrossRef](#)] [[PubMed](#)]

108. Tassi Yunga, S.; Gower, A.J.; Melrose, A.R.; Fitzgerald, M.K.; Rajendran, A.; Lusardi, T.A.; Armstrong, R.J.; Minnier, J.; Jordan, K.R.; McCarty, O.J.T.; et al. Effects of ex vivo blood anticoagulation and preanalytical processing time on the proteome content of platelets. *J. Thromb. Haemost.* **2022**, *20*, 1437–1450. [[CrossRef](#)] [[PubMed](#)]
109. Karsten, E.; Breen, E.; Herbert, B.R. Red blood cells are dynamic reservoirs of cytokines. *Sci. Rep.* **2018**, *8*, 3101. [[CrossRef](#)] [[PubMed](#)]
110. Molloy, M.P.; Hill, C.; O'Rourke, M.B.; Chandra, J.; Steffen, P.; McKay, M.J.; Pascovici, D.; Herbert, B.R. Proteomic Analysis of Whole Blood Using Volumetric Absorptive Microsampling for Precision Medicine Biomarker Studies. *J. Proteome Res.* **2022**, *21*, 1196–1203. [[CrossRef](#)]
111. Chevallet, M.; Santoni, V.; Poinas, A.; Rouquié, D.; Fuchs, A.; Kieffer, S.; Kieffer, S.; Lunardi, J.; Garin, J.; Rabilloud, T. New zwitterionic detergents improve the analysis of membrane proteins by two-dimensional electrophoresis. *Electrophoresis* **1998**, *19*, 1901–1909. [[CrossRef](#)]
112. Weiss, W.; Görg, A. Sample Solubilization Buffers for Two-Dimensional Electrophoresis. In *2D PAGE: Sample Preparation and Fractionation*; Posch, A., Ed.; Humana Press: Totowa, NJ, USA, 2008; pp. 35–42.
113. Churchward, M.A.; Butt, R.H.; Lang, J.C.; Hsu, K.K.; Coorsen, J.R. Enhanced detergent extraction for analysis of membrane proteomes by two-dimensional gel electrophoresis. *Proteome Sci.* **2005**, *3*, 5. [[CrossRef](#)] [[PubMed](#)]
114. Butt, R.H.; Coorsen, J.R. Postfractionation for Enhanced Proteomic Analyses: Routine Electrophoretic Methods Increase the Resolution of Standard 2D-PAGE. *J. Proteome Res.* **2005**, *4*, 982–991. [[CrossRef](#)] [[PubMed](#)]
115. Butt, R.H.; Coorsen, J.R. Pre-extraction sample handling by automated frozen disruption significantly improves subsequent proteomic analyses. *J. Proteome Res.* **2006**, *5*, 437–448. [[CrossRef](#)]
116. Butt, R.H.; Lee, M.W.Y.; Pirshahid, S.A.; Backlund, P.S.; Wood, S.; Coorsen, J.R. An Initial Proteomic Analysis of Human Preterm Labor: Placental Membranes. *J. Proteome Res.* **2006**, *5*, 3161–3172. [[CrossRef](#)]
117. Partridge, M.A.; Gopinath, S.; Myers, S.J.; Coorsen, J.R. An initial top-down proteomic analysis of the standard cuprizone mouse model of multiple sclerosis. *J. Chem. Biol.* **2016**, *9*, 9–18. [[CrossRef](#)]
118. Hibbert, J.E.; Butt, R.H.; Coorsen, J.R. Actin is not an essential component in the mechanism of calcium-triggered vesicle fusion. *Int. J. Biochem. Cell Biol.* **2006**, *38*, 461–471. [[CrossRef](#)] [[PubMed](#)]
119. Furber, K.L.; Backlund, P.S.; Yergey, A.L.; Coorsen, J.R. Unbiased Thiol-Labeling and Top-Down Proteomic Analyses Implicate Multiple Proteins in the Late Steps of Regulated Secretion. *Proteomes* **2019**, *7*, 34. [[CrossRef](#)] [[PubMed](#)]
120. Sen, M.K.; Almuslehi, M.S.M.; Gyengesi, E.; Myers, S.J.; Shortland, P.J.; Mahns, D.A.; Coorsen, J.R. Suppression of the Peripheral Immune System Limits the Central Immune Response Following Cuprizone-Feeding: Relevance to Modelling Multiple Sclerosis. *Cells* **2019**, *8*, 1314. [[CrossRef](#)] [[PubMed](#)]
121. Almuslehi, M.S.M.; Sen, M.K.; Shortland, P.J.; Mahns, D.A.; Coorsen, J.R. Histological and Top-Down Proteomic Analyses of the Visual Pathway in the Cuprizone Demyelination Model. *J. Mol. Neurosci.* **2022**, *72*, 1374–1401. [[CrossRef](#)] [[PubMed](#)]
122. Carbonara, K.; Padula, M.P.; Coorsen, J.R. Quantitative assessment confirms deep proteome analysis by integrative top-down proteomics. *Electrophoresis* **2022**, *44*, 472–480. [[CrossRef](#)] [[PubMed](#)]
123. Chen, E.I.; Cociorva, D.; Norris, J.L.; Yates, J.R. Optimization of Mass Spectrometry-Compatible Surfactants for Shotgun Proteomics. *J. Proteome Res.* **2007**, *6*, 2529–2538. [[CrossRef](#)] [[PubMed](#)]
124. Brown, K.A.; Chen, B.; Guardado-Alvarez, T.M.; Lin, Z.; Hwang, L.; Ayaz-Guner, S.; Jin, S.; Ge, Y. A photocleavable surfactant for top-down proteomics. *Nat. Methods* **2019**, *16*, 417–420. [[CrossRef](#)] [[PubMed](#)]
125. Goulden, T.; Bodachivskyi, I.; Padula, M.P.; Williams, D.B.G. Concentrated ionic liquids for proteomics: Caveat emptor! *Int. J. Biol. Macromol.* **2023**, *253*, 127438. [[CrossRef](#)]
126. Woodland, B.; Necakov, A.; Coorsen, J.R. Optimized Proteome Reduction for Integrative Top-Down Proteomics. *Proteomes* **2023**, *11*, 10. [[CrossRef](#)]
127. Stimpson, S.E.; Coorsen, J.R.; Myers, S.J. Optimal isolation of mitochondria for proteomic analyses. *Anal. Biochem.* **2015**, *475*, 1–3. [[CrossRef](#)]
128. D'Silva, A.M.; Hyett, J.A.; Coorsen, J.R. Proteomic analysis of first trimester maternal serum to identify candidate biomarkers potentially predictive of spontaneous preterm birth. *J. Proteom.* **2018**, *178*, 31–42. [[CrossRef](#)] [[PubMed](#)]
129. Coorsen, J.R.; Blank, P.S.; Tahara, M.; Zimmerberg, J. Biochemical and Functional Studies of Cortical Vesicle Fusion: The SNARE Complex and Ca²⁺ Sensitivity. *J. Cell Biol.* **1998**, *143*, 1845–1857. [[CrossRef](#)] [[PubMed](#)]
130. Ye, Z.; Sabatier, P.; Hoeven, L.V.D.; Phlairaharn, T.; Hartlmayr, D.; Izaguirre, F.; Seth, A.; Joshi, H.J.; Bekker-Jensen, D.B.; Bache, N.; et al. High-throughput and scalable single cell proteomics identifies over 5000 proteins per cell. *bioRxiv* **2023**. [[CrossRef](#)]
131. Rogasevskaia, T.P.; Coorsen, J.R. A new approach to the molecular analysis of docking, priming, and regulated membrane fusion. *J. Chem. Biol.* **2011**, *4*, 117–136. [[CrossRef](#)] [[PubMed](#)]
132. Noaman, N.; Coorsen, J.R. Coomassie does it (better): A Robin Hood approach to total protein quantification. *Anal. Biochem.* **2018**, *556*, 53–56. [[CrossRef](#)]
133. Görg, A.; Postel, W.; Weser, J.; Günther, S.; Strahler, J.R.; Hanash, S.M.; Somerlot, L. Horizontal two-dimensional electrophoresis with immobilized pH gradients in the first dimension in the presence of nonionic detergent. *Electrophoresis* **1987**, *8*, 45–51. [[CrossRef](#)]
134. Taylor, R.C.; Coorsen, J.R. Proteome Resolution by Two-Dimensional Gel Electrophoresis Varies with the Commercial Source of IPG Strips. *J. Proteome Res.* **2006**, *5*, 2919–2927. [[CrossRef](#)]

135. Wright, E.P.; Padula, M.P.; Higgins, V.J.; Aldrich-Wright, J.R.; Coorsen, J.R. A Systems Biology Approach to Understanding the Mechanisms of Action of an Alternative Anticancer Compound in Comparison to Cisplatin. *Proteomes* **2014**, *2*, 501–526. [[CrossRef](#)]
136. Stroud, L.J.; Slapeta, J.; Padula, M.P.; Druery, D.; Tsiotsioras, G.; Coorsen, J.R.; Stack, C.M. Comparative proteomic analysis of two pathogenic *Trichomonas foetus* genotypes: There is more to the proteome than meets the eye. *Int. J. Parasitol.* **2017**, *47*, 203–213. [[CrossRef](#)]
137. Mazinani, S.A.; Noaman, N.; Pergande, M.R.; Cologna, S.M.; Coorsen, J.; Yan, H. Exposure to microwave irradiation at constant culture temperature slows the growth of *Escherichia coli* DE3 cells, leading to modified proteomic profiles. *RSC Adv.* **2019**, *9*, 11810–11817. [[CrossRef](#)] [[PubMed](#)]
138. Gauci, V.J.; Wright, E.P.; Coorsen, J.R. Quantitative proteomics: Assessing the spectrum of in-gel protein detection methods. *J. Chem. Biol.* **2011**, *4*, 3–29. [[CrossRef](#)] [[PubMed](#)]
139. Mansour née Gauci, V.J.; Noaman, N.; Coorsen, J.R. Gel-Staining Techniques—Dyeing to Know It All. In *Encyclopedia of Life Sciences*; Wiley: Hoboken, NJ, USA, 2016; pp. 1–10.
140. Carbonara, K.; Coorsen, J.R. Sometimes faster can be better: Microneedling IPG strips enables higher throughput for integrative top-down proteomics. *Proteomics* **2022**, *23*, e2200307. [[CrossRef](#)] [[PubMed](#)]
141. Coorsen, J.R.; Blank, P.S.; Albertorio, F.; Bezrukov, L.; Kolosova, I.; Backlund, P.S.; Zimmerberg, J. Quantitative femto- to attomole immunodetection of regulated secretory vesicle proteins critical to exocytosis. *Anal. Biochem.* **2002**, *307*, 54–62. [[CrossRef](#)] [[PubMed](#)]
142. Carbonara, K.; Coorsen, J.R. A ‘green’ approach to fixing polyacrylamide gels. *Anal. Biochem.* **2020**, *605*, 113853. [[CrossRef](#)] [[PubMed](#)]
143. Harris, L.R.; Churchward, M.A.; Butt, R.H.; Coorsen, J.R. Assessing Detection Methods for Gel-Based Proteomic Analyses. *J. Proteome Res.* **2007**, *6*, 1418–1425. [[CrossRef](#)]
144. Butt, R.H.; Coorsen, J.R. Coomassie Blue as a Near-infrared Fluorescent Stain: A Systematic Comparison with Sypro Ruby for In-gel Protein Detection*. *Mol. Cell. Proteom.* **2013**, *12*, 3834–3850. [[CrossRef](#)]
145. Gauci, V.J.; Padula, M.P.; Coorsen, J.R. Coomassie blue staining for high sensitivity gel-based proteomics. *J. Proteom.* **2013**, *90*, 96–106. [[CrossRef](#)]
146. Noaman, N.; Abbineni, P.S.; Withers, M.; Coorsen, J.R. Coomassie staining provides routine (sub)femtomole in-gel detection of intact proteoforms: Expanding opportunities for genuine Top-down Proteomics. *Electrophoresis* **2017**, *38*, 3086–3099. [[CrossRef](#)]
147. Wright, E.P.; Prasad, K.A.G.; Padula, M.P.; Coorsen, J.R. Deep Imaging: How Much of the Proteome Does Current Top-Down Technology Already Resolve? *PLoS ONE* **2014**, *9*, e86058. [[CrossRef](#)]
148. Brauner, J.M.; Groemer, T.W.; Stroebel, A.; Grosse-Holz, S.; Oberstein, T.; Wiltfang, J.; Kornhuber, J.; Maler, J.M. Spot quantification in two dimensional gel electrophoresis image analysis: Comparison of different approaches and presentation of a novel compound fitting algorithm. *BMC Bioinform.* **2014**, *15*, 181. [[CrossRef](#)] [[PubMed](#)]
149. Kostopoulou, E.; Katsigiannis, S.; Maroulis, D. 2D-gel spot detection and segmentation based on modified image-aware grow-cut and regional intensity information. *Comput. Methods Programs Biomed.* **2015**, *122*, 26–39. [[CrossRef](#)] [[PubMed](#)]
150. Marczyk, M. Mixture Modeling of 2-D Gel Electrophoresis Spots Enhances the Performance of Spot Detection. *IEEE Trans. NanoBioscience* **2017**, *16*, 91–99. [[CrossRef](#)] [[PubMed](#)]
151. Molina-Mora, J.A.; Chinchilla-Montero, D.; Castro-Peña, C.; García, F. Two-dimensional gel electrophoresis (2D-GE) image analysis based on CellProfiler: *Pseudomonas aeruginosa* AG1 as model. *Medicine* **2020**, *99*, e23373. [[CrossRef](#)] [[PubMed](#)]
152. Tiwari, A.; Williams, W.P.; Shan, X. MatGel: A MATLAB program for quantitative analysis of 2D polyacrylamide electrophoresis (2D-PAGE) protein gel images. *MethodsX* **2022**, *9*, 101930. [[CrossRef](#)] [[PubMed](#)]
153. Matuzevičius, D. Synthetic Data Generation for the Development of 2D Gel Electrophoresis Protein Spot Models. *Appl. Sci.* **2022**, *12*, 4393. [[CrossRef](#)]
154. Mogi, M.; Kojima, K.; Harada, M.; Nagatsu, T. Purification and immunochemical properties of tyrosine hydroxylase in human brain. *Neurochem. Int.* **1986**, *8*, 423–428. [[CrossRef](#)] [[PubMed](#)]
155. Ehrhart, J.C.; Duthu, A.; Ullrich, S.; Appella, E.; May, P. Specific interaction between a subset of the p53 protein family and heat shock proteins hsp72/hsc73 in a human osteosarcoma cell line. *Oncogene* **1988**, *3*, 595–603. [[PubMed](#)]
156. Traub, P.; Scherbarth, A.; Willingale-Theune, J.; Traub, U. Large scale co-isolation of vimentin and nuclear lamins from ehrlich ascites tumor cells cultured in vitro. *Prep. Biochem.* **1988**, *18*, 381–404. [[CrossRef](#)]
157. Jäger, D.; Seliger, C.; Redpath, N.T.; Friedrich, I.; Silber, R.-E.; Pönicke, K.; Werdan, K.; Müller-Werdan, U. Heterogeneity of cardiac rat and human elongation factor 2. *Electrophoresis* **2000**, *21*, 2729–2736. [[CrossRef](#)]
158. Zhan, X.; Desiderio, D.M. The human pituitary nitroproteome: Detection of nitrotyrosyl-proteins with two-dimensional Western blotting, and amino acid sequence determination with mass spectrometry. *Biochem. Biophys. Res. Commun.* **2004**, *325*, 1180–1186. [[CrossRef](#)] [[PubMed](#)]
159. Guidi, F.; Magherini, F.; Gamberi, T.; Bini, L.; Puglia, M.; Marzocchini, R.; Ranaldi, F.; Modesti, P.A.; Gulisano, M.; Modesti, A. Plasma protein carbonylation and physical exercise. *Mol. BioSystems* **2011**, *7*, 640–650. [[CrossRef](#)] [[PubMed](#)]
160. Wizeman, J.W.; Nicholas, A.P.; Ishigami, A.; Mohan, R. Citrullination of glial intermediate filaments is an early response in retinal injury. *Mol. Vis.* **2016**, *22*, 1137. [[PubMed](#)]

161. Kusch, K.; Uecker, M.; Liepold, T.; Möbius, W.; Hoffmann, C.; Neumann, H.; Werner, H.B.; Jahn, O. Partial Immunoblotting of 2D-Gels: A Novel Method to Identify Post-Translationally Modified Proteins Exemplified for the Myelin Acetylome. *Proteomes* **2017**, *5*, 3. [[CrossRef](#)] [[PubMed](#)]
162. Dabral, D.; Coorssen, J.R. Combined Targeted Omic and Functional Assays Identify Phospholipases A2 that Regulate Docking/Priming in Calcium-Triggered Exocytosis. *Cells* **2019**, *8*, 303. [[CrossRef](#)] [[PubMed](#)]
163. Li, B.; Wang, X.; Yang, C.; Wen, S.; Li, J.; Li, N.; Long, Y.; Mu, Y.; Liu, J.; Liu, Q.; et al. Human growth hormone proteoform pattern changes in pituitary adenomas: Potential biomarkers for 3P medical approaches. *EPMA J.* **2021**, *12*, 67–89. [[CrossRef](#)]
164. Scheler, C.; Müller, E.; Stahl, J.; Müller-Werdan, U.; Salnikow, J.; Jungblut, P. Identification and characterization of heat shock protein 27 protein species in human myocardial two-dimensional electrophoresis patterns. *Electrophoresis* **1997**, *18*, 2823–2831. [[CrossRef](#)] [[PubMed](#)]
165. Kendrick, N.; Darie, C.C.; Hoelter, M.; Powers, G.; Johansen, J. 2D SDS PAGE in Combination with Western Blotting and Mass Spectrometry Is a Robust Method for Protein Analysis with Many Applications. In *Advancements of Mass Spectrometry in Biomedical Research*; Woods, A.G., Darie, C.C., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 563–574.
166. Jones, K.S.; Chapman, A.E.; Driscoll, H.A.; Fuller, E.P.; Kelly, M.; Li, X.; Mansour, S.; McBride, S.L.; Zhao, Q.; Weiner, M.; et al. MILKSHAKE: Novel validation method for antibodies to post-translationally modified targets by surrogate Western blot. *Biotechniques* **2022**, *72*, 11–20. [[CrossRef](#)]
167. Coorssen, J.R.; Blank, P.S.; Albertorio, F.; Bezrukov, L.; Kolosova, I.; Chen, X.; Backlund, P.S., Jr.; Zimmerberg, J. Regulated secretion: SNARE density, vesicle fusion and calcium dependence. *J. Cell Sci.* **2003**, *116*, 2087–2097. [[CrossRef](#)]
168. Kurgan, N.; Noaman, N.; Pergande, M.R.; Cologna, S.M.; Coorssen, J.R.; Klentrou, P. Changes to the Human Serum Proteome in Response to High Intensity Interval Exercise: A Sequential Top-Down Proteomic Analysis. *Front. Physiol.* **2019**, *10*, 362. [[CrossRef](#)]
169. Hu, M.; Liu, Y.; Yu, K.; Liu, X. Decreasing the amount of trypsin in in-gel digestion leads to diminished chemical noise and improved protein identifications. *J. Proteom.* **2014**, *109*, 16–25. [[CrossRef](#)] [[PubMed](#)]
170. Mansuri, M.S.; Bathla, S.; Lam, T.T.; Nairn, A.C.; Williams, K.R. Optimal conditions for carrying out trypsin digestions on complex proteomes: From bulk samples to single cells. *J. Proteom.* **2024**, *297*, 105109. [[CrossRef](#)] [[PubMed](#)]
171. Woessmann, J.A.-O.; Petrosius, V.; Üresin, N.; Kotol, D.A.-O.; Aragon-Fernandez, P.; Hober, A.A.-O.; Auf dem Keller, U.A.-O.; Edfors, F.; Schoof, E.A.-O. Assessing the Role of Trypsin in Quantitative Plasma and Single-Cell Proteomics toward Clinical Application. *Anal. Chem.* **2023**, *95*, 13649–13658. [[CrossRef](#)] [[PubMed](#)]
172. Zheng, Y.Z.; DeMarco, M.L. Manipulating trypsin digestion conditions to accelerate proteolysis and simplify digestion workflows in development of protein mass spectrometric assays for the clinical laboratory. *Clin. Mass Spectrom.* **2017**, *6*, 1–12. [[CrossRef](#)]
173. Shuford, C.M.; Grant, R.P. Cheaper, faster, simpler trypsin digestion for high-throughput targeted protein quantification. *J. Mass Spectrom. Adv. Clin. Lab.* **2023**, *30*, 74–82. [[CrossRef](#)] [[PubMed](#)]
174. Wei, X.; Liu, P.N.; Mooney, B.P.; Nguyen, T.T.; Greenlief, C.M. A Comprehensive Study of Gradient Conditions for Deep Proteome Discovery in a Complex Protein Matrix. *Int. J. Mol. Sci.* **2022**, *23*, 11714. [[CrossRef](#)] [[PubMed](#)]
175. Jiang, Y.; Rex, D.A.B.; Schuster, D.; Neely, B.A.; Rosano, G.L.; Volkmar, N.; Momenzadeh, A.; Peters-Clarke, T.M.; Egbert, S.B.; Kreimer, S.; et al. Comprehensive Overview of Bottom-Up Proteomics using Mass Spectrometry. *arXiv* **2023**, arXiv:2311.07791v1.
176. Houel, S.; Abernathy, R.; Renganathan, K.; Meyer-Arendt, K.; Ahn, N.G.; Old, W.M. Quantifying the impact of chimera MS/MS spectra on peptide identification in large-scale proteomics studies. *J. Proteome Res.* **2010**, *9*, 4152–4160. [[CrossRef](#)] [[PubMed](#)]
177. Adair, L.R.; Jones, I.; Cramer, R. Utilizing Precursor Ion Connectivity of Different Charge States to Improve Peptide and Protein Identification in MS/MS Analysis. *Anal. Chem.* **2024**, *96*, 985–990. [[CrossRef](#)]
178. Frankenfield, A.M.; Ni, J.; Ahmed, M.; Hao, L. Protein Contaminants Matter: Building Universal Protein Contaminant Libraries for DDA and DIA Proteomics. *J. Proteome Res.* **2022**, *21*, 2104–2113. [[CrossRef](#)]
179. Hirabayashi, A.; Ishimaru, M.; Manri, N.; Yokosuka, T.; Hanzawa, H. Detection of potential ion suppression for peptide analysis in nanoflow liquid chromatography/mass spectrometry. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 2860–2866. [[CrossRef](#)]
180. Cramer, R. High-speed Analysis of Large Sample Sets—How Can This Key Aspect of the Omics Be Achieved? *Mol. Cell Proteom.* **2020**, *19*, 1760–1766. [[CrossRef](#)]
181. Slavov, N. Increasing proteomics throughput. *Nat. Biotechnol.* **2021**, *39*, 809–810. [[CrossRef](#)]
182. Piehowski, P.D.; Petyuk, V.A.; Orton, D.J.; Xie, F.; Moore, R.J.; Ramirez-Restrepo, M.; Engel, A.; Lieberman, A.P.; Albin, R.L.; Camp, D.G.; et al. Sources of technical variability in quantitative LC-MS proteomics: Human brain tissue sample analysis. *J. Proteome Res.* **2013**, *12*, 2128–2137. [[CrossRef](#)]
183. Bian, Y.; Zheng, R.; Bayer, F.P.; Wong, C.; Chang, Y.-C.; Meng, C.; Zolg, D.P.; Reinecke, M.; Zecha, J.; Wiechmann, S.; et al. Robust, reproducible and quantitative analysis of thousands of proteomes by micro-flow LC-MS/MS. *Nat. Commun.* **2020**, *11*, 157. [[CrossRef](#)]
184. Compton, P.D.; Zamdborg, L.; Thomas, P.M.; Kelleher, N.L. On the Scalability and Requirements of Whole Protein Mass Spectrometry. *Anal. Chem.* **2011**, *83*, 6868–6874. [[CrossRef](#)]
185. Meyer, S.; Clases, D.; Gonzalez de Vega, R.; Padula, M.P.; Doble, P.A. Separation of intact proteins by capillary electrophoresis. *Analyst* **2022**, *147*, 2988–2996. [[CrossRef](#)]
186. Schwenzer, A.K.; Kruse, L.; Jooss, K.; Neuss, C. Capillary electrophoresis-mass spectrometry for protein analyses under native conditions: Current progress and perspectives. *Proteomics* **2023**, *24*, e2300135. [[CrossRef](#)]

187. Ridgeway, M.E.; Lubeck, M.; Jordens, J.; Mann, M.; Park, M.A. Trapped ion mobility spectrometry: A short review. *Int. J. Mass Spectrom.* **2018**, *425*, 22–35. [[CrossRef](#)]
188. Meier, F.; Park, M.A.; Mann, M. Trapped Ion Mobility Spectrometry and Parallel Accumulation Serial Fragmentation in Proteomics. *Mol. Cell. Proteom.* **2021**, *20*, 100138. [[CrossRef](#)]
189. Meier, F.; Brunner, A.D.; Koch, S.; Koch, H.; Lubeck, M.; Krause, M.; Goedecke, N.; Decker, J.; Kosinski, T.; Park, M.A.; et al. Online Parallel Accumulation-Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer. *Mol. Cell Proteom.* **2018**, *17*, 2534–2545. [[CrossRef](#)]
190. Kafader, J.O.; Durbin, K.R.; Melani, R.D.; Des Soye, B.J.; Schachner, L.F.; Senko, M.W.; Compton, P.D.; Kelleher, N.L. Individual Ion Mass Spectrometry Enhances the Sensitivity and Sequence Coverage of Top-Down Mass Spectrometry. *J. Proteome Res.* **2020**, *19*, 1346–1350. [[CrossRef](#)]
191. Tyanova, S.; Temu, T.; Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **2016**, *11*, 2301–2319. [[CrossRef](#)]
192. Shah, A.D.; Goode, R.J.A.; Huang, C.; Powell, D.R.; Schittenhelm, R.B. LFQ-Analyst: An Easy-To-Use Interactive Web Platform to Analyze and Visualize Label-Free Proteomics Data Preprocessed with MaxQuant. *J. Proteome Res.* **2020**, *19*, 204–211. [[CrossRef](#)]
193. Yu, F.; Teo, G.C.; Kong, A.T.; Haynes, S.E.; Avtonomov, D.M.; Geiszler, D.J.; Nesvizhskii, A.I. Identification of modified peptides using localization-aware open search. *Nat. Commun.* **2020**, *11*, 4065. [[CrossRef](#)]
194. Yi, H.; Haijian, Z.; Ginny Xiaohu, L.; Yamei, D.; Fengchao, Y.; Hossein Valipour, K.; Joel, R.S.; Ralf, B.S.; Alexey, I.N. Analysis and visualization of quantitative proteomics data using FragPipe-Analyst. *bioRxiv* **2024**. [[CrossRef](#)]
195. Kitata, R.B.; Yang, J.C.; Chen, Y.J. Advances in data-independent acquisition mass spectrometry towards comprehensive digital proteome landscape. *Mass Spectrom. Rev.* **2023**, *42*, 2324–2348. [[CrossRef](#)]
196. Geromanos, S.J.; Vissers, J.P.; Silva, J.C.; Dorschel, C.A.; Li, G.Z.; Gorenstein, M.V.; Bateman, R.H.; Langridge, J.I. The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS. *Proteomics* **2009**, *9*, 1683–1695. [[CrossRef](#)]
197. Demichev, V.A.-O.; Messner, C.B.; Vernardis, S.I.; Lilley, K.A.-O.; Ralser, M.A.-O. DIA-NN: Neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **2020**, *17*, 41–44. [[CrossRef](#)]
198. Valentine, S.J.; Ewing, M.A.; Dilger, J.M.; Glover, M.S.; Geromanos, S.; Hughes, C.; Clemmer, D.E. Using ion mobility data to improve peptide identification: Intrinsic amino acid size parameters. *J. Proteome Res.* **2011**, *10*, 2318–2329. [[CrossRef](#)]
199. Vitko, D.; Chou, W.-F.; Nouri Golmaei, S.; Lee, J.-Y.; Belthangady, C.; Blume, J.; Chan, J.K.; Flores-Campuzano, G.; Hu, Y.; Liu, M.; et al. timsTOF HT Improves Protein Identification and Quantitative Reproducibility for Deep Unbiased Plasma Protein Biomarker Discovery. *J. Proteome Res.* **2024**, *23*, 929–938. [[CrossRef](#)]
200. Demichev, V.; Szyrwiell, L.; Yu, F.; Teo, G.C.; Rosenberger, G.; Niewianda, A.; Ludwig, D.; Decker, J.; Kaspar-Schoenefeld, S.; Lilley, K.S.; et al. dia-PASEF data analysis using FragPipe and DIA-NN for deep proteomics of low sample amounts. *Nat. Commun.* **2022**, *13*, 3944. [[CrossRef](#)]
201. Chang, C.-H.; Ishihama, Y. Deep profiling of proteomics dataset by liquid chromatography/ trapped ion mobility spectrometry/tandem mass spectrometry. *J. Proteome Data Methods* **2022**, *4*, 3. [[CrossRef](#)]
202. Ahmad, R.; Budnik, B. A review of the current state of single-cell proteomics and future perspective. *Anal. Bioanal. Chem.* **2023**, *415*, 6889–6899. [[CrossRef](#)]
203. Rosenberger, F.A.; Thielert, M.; Strauss, M.T.; Schweizer, L.; Ammar, C.; Mädler, S.C.; Metousis, A.; Skowronek, P.; Wahle, M.; Madden, K.; et al. Spatial single-cell mass spectrometry defines zonation of the hepatocyte proteome. *Nat. Methods* **2023**, *20*, 1530–1536. [[CrossRef](#)]
204. Zhang, Y.; Sohn, C.; Lee, S.; Ahn, H.; Seo, J.; Cao, J.; Cai, L. Detecting protein and post-translational modifications in single cells with iDentification and qUantification sEparatiOn (DUET). *Commun. Biol.* **2020**, *3*, 420. [[CrossRef](#)]
205. Orsburn, B.C.; Yuan, Y.; Bumpus, N.N. Insights into protein post-translational modification landscapes of individual human cells by trapped ion mobility time-of-flight mass spectrometry. *Nat. Commun.* **2022**, *13*, 7246. [[CrossRef](#)]
206. Huffman, R.G.; Leduc, A.; Wichmann, C.; Di Gioia, M.; Borriello, F.; Specht, H.; Derks, J.; Khan, S.; Khoury, L.; Emmott, E.; et al. Prioritized mass spectrometry increases the depth, sensitivity and data completeness of single-cell proteomics. *Nat. Methods* **2023**, *20*, 714–722. [[CrossRef](#)]
207. Melby, J.A.; Brown, K.A.; Gregorich, Z.R.; Roberts, D.S.; Chapman, E.A.; Ehlers, L.E.; Gao, Z.; Larson, E.J.; Jin, Y.; Lopez, J.R.; et al. High sensitivity top-down proteomics captures single muscle cell heterogeneity in large proteoforms. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2222081120. [[CrossRef](#)]
208. Mund, A.; Coscia, F.; Kriston, A.; Hollandi, R.; Kovács, F.; Brunner, A.-D.; Migh, E.; Schweizer, L.; Santos, A.; Bzorek, M.; et al. Deep Visual Proteomics defines single-cell identity and heterogeneity. *Nat. Biotechnol.* **2022**, *40*, 1231–1240. [[CrossRef](#)]
209. Sutandy, F.X.; Qian, J.; Chen, C.S.; Zhu, H. Overview of protein microarrays. *Curr. Protoc. Protein Sci.* **2013**, *72*, 27.1.1–27.1.16. [[CrossRef](#)]
210. Aparna, G.M.; Tetala, K.K.R. Recent Progress in Development and Application of DNA, Protein, Peptide, Glycan, Antibody, and Aptamer Microarrays. *Biomolecules* **2023**, *13*, 602. [[CrossRef](#)]
211. Darmanis, S.; Nong, R.Y.; Vänelid, J.; Siegbahn, A.; Ericsson, O.; Fredriksson, S.; Bäcklin, C.; Gut, M.; Heath, S.; Gut, I.G.; et al. ProteinSeq: High-Performance Proteomic Analyses by Proximity Ligation and Next Generation Sequencing. *PLoS ONE* **2011**, *6*, e25583. [[CrossRef](#)]

212. Nong, R.Y.; Wu, D.; Yan, J.; Hammond, M.; Gu, G.J.; Kamali-Moghaddam, M.; Landegren, U.; Darmanis, S. Solid-phase proximity ligation assays for individual or parallel protein analyses with readout via real-time PCR or sequencing. *Nat. Protoc.* **2013**, *8*, 1234–1248. [CrossRef]
213. Gold, L.; Walker, J.J.; Wilcox, S.K.; Williams, S. Advances in human proteomics at high scale with the SOMAscan proteomics platform. *New Biotechnol.* **2012**, *29*, 543–549. [CrossRef]
214. Zhang, M.; Tang, C.; Wang, Z.; Chen, S.; Zhang, D.; Li, K.; Sun, K.; Zhao, C.; Wang, Y.; Xu, M.; et al. Real-time detection of 20 amino acids and discrimination of pathologically relevant peptides with functionalized nanopore. *Nat. Methods* **2024**, *21*, 609–618. [CrossRef]
215. Yu, L.; Kang, X.; Li, F.; Mehrafrooz, B.; Makhamreh, A.; Fallahi, A.; Foster, J.C.; Aksimentiev, A.; Chen, M.; Wanunu, M. Unidirectional single-file transport of full-length proteins through a nanopore. *Nat. Biotechnol.* **2023**, *41*, 1130–1139. [CrossRef]
216. Nova, I.C.; Ritmejeris, J.; Brinkerhoff, H.; Koenig, T.J.R.; Gundlach, J.H.; Dekker, C. Detection of phosphorylation post-translational modifications along single peptides with nanopores. *Nat. Biotechnol.* **2023**, 1–5. [CrossRef]
217. Martin-Baniandres, P.; Lan, W.-H.; Board, S.; Romero-Ruiz, M.; Garcia-Manyes, S.; Qing, Y.; Bayley, H. Enzyme-less nanopore detection of post-translational modifications within long polypeptides. *Nat. Nanotechnol.* **2023**, *18*, 1335–1340. [CrossRef]
218. Filius, M.; van Wee, R.; de Lannoy, C.; Westerlaken, I.; Li, Z.; Kim, S.H.; de Agrela Pinto, C.; Wu, Y.; Boons, G.-J.; Pabst, M.; et al. Full-length single-molecule protein fingerprinting. *Nat. Nanotechnol.* **2024**, 1–8. [CrossRef] [PubMed]
219. von Mering, C.; Huynen, M.; Jaeggi, D.; Schmidt, S.; Bork, P.; Snel, B. STRING: A database of predicted functional associations between proteins. *Nucleic Acids Res.* **2003**, *31*, 258–261. [CrossRef]
220. Szklarczyk, D.; Kirsch, R.; Koutrouli, M.; Nastou, K.; Mehryary, F.; Hachilif, R.; Gable, A.L.; Fang, T.; Doncheva, N.T.; Pyysalo, S.; et al. The STRING database in 2023: Protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **2023**, *51*, D638–D646. [CrossRef] [PubMed]
221. Mi, H.; Muruganujan, A.; Huang, X.; Ebert, D.; Mills, C.; Guo, X.; Thomas, P.D. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat. Protoc.* **2019**, *14*, 703–721. [CrossRef]
222. Gillespie, M.; Jassal, B.; Stephan, R.; Milacic, M.; Rothfels, K.; Senff-Ribeiro, A.; Griss, J.; Sevilla, C.; Matthews, L.; Gong, C.; et al. The reactome pathway knowledgebase 2022. *Nucleic Acids Res.* **2022**, *50*, D687–D692. [CrossRef] [PubMed]
223. LeDuc, R.D.; Schwämmle, V.; Shortreed, M.R.; Cesnik, A.J.; Solntsev, S.K.; Shaw, J.B.; Martin, M.J.; Vizcaino, J.A.; Alpi, E.; Danis, P.; et al. ProForma: A Standard Proteoform Notation. *J. Proteome Res.* **2018**, *17*, 1321–1325. [CrossRef] [PubMed]
224. Huang, C.-F.; Kline, J.T.; Negrão, F.; Robey, M.T.; Toby, T.K.; Durbin, K.R.; Fellers, R.T.; Friedewald, J.J.; Levitsky, J.; Abecassis, M.M.I.; et al. Targeted Quantification of Proteoforms in Complex Samples by Proteoform Reaction Monitoring. *Anal. Chem.* **2024**, *96*, 3578–3586. [CrossRef] [PubMed]
225. Terwilliger, T.C.; Liebschner, D.; Croll, T.I.; Williams, C.J.; McCoy, A.J.; Poon, B.K.; Afonine, P.V.; Oeffner, R.D.; Richardson, J.S.; Read, R.J.; et al. AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination. *Nat. Methods* **2024**, *21*, 110–116. [CrossRef] [PubMed]
226. Cheng, J.; Novati, G.; Pan, J.; Bycroft, C.; Žemgulytė, A.; Applebaum, T.; Pritzel, A.; Wong, L.H.; Zielinski, M.; Sargeant, T.; et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* **2023**, *381*, eadg7492. [CrossRef]
227. Nagasawa, I.; Muroi, M.; Kawatani, M.; Ohishi, T.; Ohba, S.I.; Kawada, M.; Osada, H. Identification of a Small Compound Targeting PKM2-Regulated Signaling Using 2D Gel Electrophoresis-Based Proteome-wide CETSA. *Cell Chem. Biol.* **2020**, *27*, 186–196.e184. [CrossRef]
228. Artificial Intelligence Is Taking Over Drug Development. Available online: <https://www.economist.com/technology-quarterly/2024/03/27/artificial-intelligence-is-taking-over-drug-development> (accessed on 16 April 2024).
229. Will, A.; Oliinyk, D.; Bleiholder, C.; Meier, F. Peptide collision cross sections of 22 post-translational modifications. *Anal. Bioanal. Chem.* **2023**, *415*, 6633–6645. [CrossRef]
230. Zecha, J.; Bayer, F.P.; Wiechmann, S.; Woortman, J.; Berner, N.; Müller, J.; Schneider, A.; Kramer, K.; Abril-Gil, M.; Hopf, T.; et al. Decrypting drug actions and protein modifications by dose- and time-resolved proteomics. *Science* **2023**, *380*, 93–101. [CrossRef]
231. Leutert, M.; Entwisle, S.W.; Villén, J. Decoding Post-Translational Modification Crosstalk with Proteomics. *Mol. Cell. Proteom.* **2021**, *20*, 100129. [CrossRef]
232. Mann, M.; Kumar, C.; Zeng, W.F.; Strauss, M.T. Artificial intelligence for proteomics and biomarker discovery. *Cell Syst.* **2021**, *12*, 759–770. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.