METHOD ARTICLE

REVISED Amicon-adapted enhanced FASP: an in-solution digestion-based alternative sample preparation method to FASP [version 2; peer review: 2 approved with reservations]David Pellerin¹, Hugo Gagnon², Jean Dubé¹, Francois Corbin¹¹Department of Biochemistry, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Sherbrooke, QC, J1H 5N4, Canada²PhenoSwitch Bioscience Inc., Sherbrooke, QC, J1H 5N4, Canada**v2** First published: 04 Jun 2015, 4:140
<https://doi.org/10.12688/f1000research.6529.1>Latest published: 24 Sep 2015, 4:140
<https://doi.org/10.12688/f1000research.6529.2>**Abstract**

Sample preparation is a crucial step for liquid chromatography-tandem mass spectrometry (LC-MS/MS)-based proteomics. Sodium dodecyl sulfate (SDS) is a powerful denaturing detergent that allows for long-term preservation of protein integrity. However, as it inhibits trypsin and interferes with LC-MS/MS analyses, it must be removed from samples prior to these experiments. The Filter-Aided Sample Preparation (FASP) method is actually one of the preferred and simplest methods for such purpose. Nonetheless, there exist great disparities in the quality of outcomes when comparing FASP to other protocols depending on the authors, and recent reports have pointed to concerns regarding its depth of proteome coverage. To address these issues, we propose an Amicon-adapted in-solution-based enhanced FASP (eFASP) approach that relies on current best practices in comprehensive proteomics sample preparation. Human megakaryoblastic leukaemia cancer cells' protein extracts were treated in parallel with both Amicon-adapted eFASP and FASP, quantified for remaining SDS and then analyzed with a 1-hr gradient LC-MS/MS run. The Amicon-adapted eFASP utilizes a passivated low molecular weight cut-off Amicon filter, and incorporates a cleaning step with a high-content deoxycholate buffer and a 'one-step-two-enzymes' trypsin/Lys-C in-solution digestion. Amicon-adapted eFASP was found more reproducible and deepened proteome coverage, especially for membrane proteins. As compared to FASP, Amicon-adapted eFASP removed much of SDS from high-protein samples and reached a notable depth of proteome coverage with nearly 1,700 proteins identified in a 1 hr LC-MS/MS single-run analysis without prior fractionation. Amicon-adapted eFASP can therefore be regarded as a simple and reliable sample preparation approach for comprehensive proteomics.

Open Peer Review**Approval Status** ? ?

	1	2
version 2 (revision) 24 Sep 2015	 view	 view
version 1 04 Jun 2015	 view	

1. **Lloyd D. Fricker**, Albert Einstein College of Medicine of Yeshiva University, Bronx, USA

2. **Éric Bonneil**, University of Montreal, Montréal, Canada

Any reports and responses or comments on the article can be found at the end of the article.

Keywords

bottom-up proteomics, Amicon, eFASP/FASP, biomarker, MEG-01, sample preparation, tandem mass spectrometry, in-solution digestion

Corresponding author: Francois Corbin (francois.corbin@usherbrooke.ca)

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by Université de Sherbrooke and the CHUS Research Center. DP holds a Frederick Banting and Charles Best Canada Graduate Scholarship Master Award from the Canadian Institutes of Health Research (CIHR) and an M.D.-M.Sc. Graduate Scholarship Master Award from the Fonds de Recherche du Québec – Santé (FRQS). HG holds a MITACS accelerated postdoctoral scholarship.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2015 Pellerin D *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

How to cite this article: Pellerin D, Gagnon H, Dubé J and Corbin F. **Amicon-adapted enhanced FASP: an in-solution digestion-based alternative sample preparation method to FASP [version 2; peer review: 2 approved with reservations]** F1000Research 2015, 4:140 <https://doi.org/10.12688/f1000research.6529.2>

First published: 04 Jun 2015, 4:140 <https://doi.org/10.12688/f1000research.6529.1>

REVISED Amendments from Version 1

We added two references (Tang *et al.* [2008], and Serang and Noble [2012]) that provide an in-depth explanation of the principles underlying the False Discovery Rate (FDR) analysis of protein identification, as well as a review of the different statistical methods for protein identification. These references clarify the ambiguity related to the proteins listed as identified with "0 peptide", which are, in fact, included in the false discovery rate inherent to our analysis. We have also re-organized our datasets according to the reviewer's comments.

See referee reports

Introduction

Proteomics plays an increasingly greater role in diverse clinical settings provided that proteins give an integrative picture of patients' phenotype and an accurate representation of changes in the status of an organism¹. As such, intense efforts are now directed toward proteomics-based biomarker discovery, disease screening, and medical diagnosis². Accuracy and reliability of bottom-up mass spectrometry (MS)-based proteomics analyses are undoubtedly dependent on the quality of prior sample preparation, which remains very challenging³. In fact, sodium dodecyl sulfate (SDS) is a powerful anionic detergent and one of the most widely used reagents for solubilisation and denaturation of proteins⁴. By binding amino acids through hydrophobic and ionic interactions, SDS alters proteins' spatial conformational structure and inhibits proteases activity, thus enabling long-term conservation of structurally preserved proteins⁵. In spite of all its advantages, SDS significantly suppresses analyte ion signals in electrospray ionization (ESI)-MS, alters the chromatographic separation of peptides during liquid chromatography (LC)^{4,6}, and strongly inhibits trypsin activity⁷, even at very low concentrations, which considerably limits protein identification. Samples therefore require to be completely depleted of SDS prior to digestion and liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. However, removing SDS while achieving a high recovery rate is a challenging task and several approaches have been proposed over the years. Methods such as gel electrophoresis separation along with in-gel digestion, protein precipitation and dialysis have shown variable efficacy, low reproducibility and limited routine applicability. Among the different protein preparation techniques recently developed, spin filter-based sample preparation, a concept first introduced by Manza *et al.*⁸ and further optimized for better detergent removal by Wisniewski *et al.*⁹ – *Filter-Aided Sample Preparation* (FASP) –, emerged as a promising approach and has already been successfully applied in diverse experimental settings^{9–12}. In fact, FASP is among the simplest sample preparation methods for SDS-containing extracts. Operating in an ultrafiltration spin filter device, the method allows samples to be cleaned of SDS with urea prior to on-filter tryptic digestion¹³. Moreover, urea enables usage of filter membrane with a molecular weight cut-off of 30 kDa (30k MWCO), while not compromising depth of proteome coverage, as it ensures unfolding of proteins during the workflow¹⁴. FASP is also described as a universal sample preparation method as it allows for unbiased proteome representation.

During past years, some groups have raised concerns regarding the depth of proteome coverage achieved with FASP^{15–20}. Furthermore, there exist great disparities in the quality of outcomes when comparing FASP to other protocols depending on the authors and also as noted in practice by laboratories attempting to use FASP. Actually, Liebler and Ham²¹ first pointed out the difficulty of removing SDS by filter-aided buffer exchange, an observation further supported by Hustoft *et al.*¹⁶. The uncharged chaotrope urea may in fact have a limited capacity to deplete SDS from proteins as it may not completely dissociate negatively charged sulfate heads from basic amino acids. Disruption of these powerful ionic interactions may in fact require substitution of SDS by another anionic surfactant, such as sodium deoxycholate (NaDoc). Moreover, Proc *et al.*²² have shown that NaDoc was the most effective detergent to deepen proteome coverage and enhance digestion reproducibility when included in the digestion buffer. As long as its concentration remains below 1%, NaDoc promotes trypsin activity by easing contact between enzyme and cleavage sites^{23–26}. This acid-insoluble surfactant is also MS-compatible as it is readily removed following acidification by either precipitation or phase transfer in organic solvent such as ethyl acetate^{24,25}. Among the other factors possibly explaining the difficulties reported with FASP is the on-filter digestion, as it may have a more limited efficacy compared to the traditional in-solution approach^{27,28}. These issues emphasize the need to fine-tune the FASP protocol with current state-of-the-art techniques in comprehensive proteomics in order to improve its robustness and reliability. As such, Erde *et al.* (2014) have recently proposed an enhanced FASP (eFASP) workflow to improve proteome coverage of SDS-solubilized protein samples¹⁸. This protocol, which is currently the only other equivalent alternative to FASP for preparing SDS-denatured proteins, incorporates a passivation step of the Microcon filter device, deoxycholic acid-based buffer exchanges, and a post-digestion phase transfer cleaning of peptides prior to LC-MS/MS analysis. These modifications led to a significant 14-fold increase in proteome coverage as compared to FASP¹⁸. Although the FASP was designed for use with the flat bottom filter Microcon device (EMD Millipore), several laboratories instead use the Amicon unit with V-shaped oriented filters (EMD Millipore)^{15,16,19,29}. Difference between filters may in part explain why Nel *et al.* (2015) failed to show a significant difference in proteome coverage between eFASP and FASP while using Amicon devices²⁹. In light of the striking discrepancy of proteome coverage between these two studies, an enhanced FASP workflow specifically tailored for Amicon filters was required in order to allow more laboratories to benefit from FASP advantages and ease-of-use in proteomics.

Herein, we report a reproducible Amicon-adapted enhanced FASP (eFASP) workflow for comprehensive proteomics that is based on the best practices in sample preparation of SDS-containing protein extracts. The Amicon-adapted eFASP method presents key modifications to the FASP protocol and addresses the above-raised concerns. We compared both methods for SDS removal efficacy, depth of proteome coverage and reproducibility. Our results show that Amicon-adapted eFASP compares favourably to FASP and can therefore be regarded as a new alternative sample preparation approach for proteomics.

Experimental section

Chemicals, reagents and materials

Ammonium bicarbonate (ABC), 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate (CHAPS), ethyl acetate, iodoacetamide, methanol, brilliant blue R250, phenylmethanesulfonyl fluoride (PMSF), potassium chloride (KCl), sodium chloride (NaCl), sodium deoxycholate (NaDoc), sodium dodecyl sulfate (SDS), stains-all, trizma base and urea were purchased from Sigma-Aldrich (St. Louis, MO) and were of the highest quality available. DL-Dithiothreitol (DTT), formamide, isopropanol, sodium hydroxide (NaOH), thiourea and Tween-20 were obtained from Thermo Fisher BioReagents (Ottawa, ON). Optima™ LC/MS grade formic acid (FA) and water were obtained from Thermo Fisher chemicals (Ottawa, ON). Unless indicated, all other chemicals were from Sigma-Aldrich (St. Louis, MO) and minimally of ACS grade. RPMI-1640 (#350–007-CL), heat-inactivated fetal bovine serum (FBS), L-glutamine and penicillin/streptomycin were obtained from Wisent Bioproducts (St-Bruno, QC). Amicon Ultra-0.5 ml centrifugal filters of 30 kDa (#UFC503096) and 10 kDa (#UFC501096) molecular weight cutoff (MWCO) were obtained from EMD Millipore Corporation (Billerica, MA). Trypsin Gold (#V5280) and Trypsin/Lys-C Mix (#V5073), both of Mass Spectrometry Grade, were from Promega (Madison, WI). BCA protein assay kit (#23227) and NanoDrop 2000c (#ND-2000c) were purchased from Thermo Scientific Pierce (Rockford, IL). Eppendorf Protein LoBind tubes were from Eppendorf (Hamburg, Germany).

Preparation of human megakaryoblastic leukaemia (MEG-01) whole cell lysates

MEG-01 cells (CRL-2021, ATCC, Manassas, VA) were grown (37°C/5% CO₂) in a medium containing 90% RPMI-1640, 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin. Adherent cells were scraped into the medium and the resulting suspension was centrifuged at 430 g (5 min/4°C) for pelleting cells. Ice-cold 0.9% NaCl-resuspended cells were again pelleted by centrifugation. Cells were then lysed with a buffer consisting of 2% SDS, 0.1 M Tris-HCl pH 8 and 10 mM DTT, and heated for 7 min at 98°C. Nucleic acids were sheared by probe sonication (Misonix Sonicator 3000, Cole-Parmer, Hills, IL). Unbroken cells and debris were cleared by centrifugation at 16,000 g for 10 min. Sonication and centrifugation were repeated once. The final supernatant was carefully withdrawn, transferred in a LoBind tube and its protein concentration quantified by bicinchoninic acid (BCA) protein assay according to manufacturer's instructions. "Non-SDS" control lysates were prepared with a buffer containing 7 M urea, 2 M thiourea, 4% CHAPS, 1% DTT and 0.5 mM PMSF. Lysates were centrifuged at 16,000 g for 20 min (4°C). Supernatants were removed and transferred in LoBind tubes. Disulphide bridges were reduced following a 1h incubation at 37°C and protein concentration quantified by the BCA protein assay.

Effect of SDS on digestion efficacy

Impact of increasing SDS concentrations (0.01% to 0.30% w/v) on trypsin (FASP) and one-step trypsin/Lys-C (Amicon-adapted eFASP) overnight digestion completion was assessed by Coomassie-stained 12% and 18% SDS-polyacrylamide gel electrophoresis (PAGE). SDS-free cell lysates were prepared by probe sonication of 50 mM

ABC-resuspended MEG-01 pellet followed by 16,000 g centrifugation (10 min). Sonication and centrifugation were repeated once on collected supernatant. Lysates were solubilized in 6 M urea, reduced for 60 min at 37°C with 10 mM DTT and alkylated by 30 mM iodoacetamide for 30 min. One hundred-microgram whole cell lysate aliquots were next diluted six-fold in ABC buffer and digested in increasing concentrations of SDS⁵. Half of samples were afterward incubated for 12h at 37°C in the presence of (1) trypsin (enzyme-to-protein ratio 1:100 w/w) or (2) trypsin/Lys-C Mix (ratio 1:25 w/w) for proteolysis. Digested samples were next mixed with SDS loading buffer, heated and half the volume loaded onto a 12% SDS-PAGE while the other half was loaded on an 18% SDS-PAGE. Gels were subsequently stained for 3h in a Coomassie solution and scanned with an Odyssey® Infrared Imaging System (LI-COR Biosciences, Lincoln, NE).

Spectrophotometric quantification of SDS

The SDS concentration was measured according to Rusconi *et al.*³⁰ by recording the absorbance (438 nm) (Ultrospec 2100 pro, Biochrom, Holliston, MA) of a stains-all dosage solution mixed with 1 µl of sample ([Supplementary Table 1](#)). To assess the assay's specificity, we performed interference studies with all solutions and SDS-free biological samples used in FASP and Amicon-adapted eFASP protocols. The complete list of tested compounds is presented in [Supplementary Table 2](#).

Filter-based sample preparation: FASP and Amicon-adapted eFASP

Refer to [Figure 1](#) for a detailed study flowchart (see also [Supplementary Table 3](#)).

Filter prerinsing

Prior to protein purification, every non-passivated filter unit has been thoroughly rinsed following manufacturer's instructions. A first cleaning spin (14,000 g/15 min [30k]/25 min [10k]) with 0.05 M NaOH was followed by a final spin with 500 µl of MS-grade H₂O.

Filter-Aided Sample Preparation (FASP)

Thirty-microliter aliquots (150 µg proteins) of MEG-01 total extract were purified by the FASP protocol⁹. Aliquots were either dispensed to prerinsed Amicon Ultra-0.5 ml 10k filters (A10k) or Amicon Ultra-0.5 ml 30k filters (A30k). Every experiment was performed in triplicate (n=3). The FASP protocol was followed exactly as described by Wisniewski *et al.*⁹. Briefly, samples were washed twice with 200 µl of 8 M urea in 0.1 M Tris-HCl pH 8.5 and spun at 14,000 g for 15 min (30k) or 30 min (10k). Flow-through was discarded from collection tube and 100 µl of 50 mM iodoacetamide solution in 8 M urea, 0.1 M Tris-HCl pH 8.5 was added to each filter unit. After mixing filter content well for one minute, cysteines were alkylated at room temperature in the dark for 20 min without shaking. Excess iodoacetamide was then removed by spinning at 14,000 g for 10 min (30k) or 30 min (10k). Samples were further washed thrice with 100 µl of 8 M urea in 0.1 M Tris-HCl pH 8.5 and centrifugation resumed each time for 15 min (30k) or 30 min (10k) at 14,000 g. Three final buffer exchanges were performed by adding 100 µl 50 mM ABC followed by a 14,000 g spin for 10 min (30k) or 30 min (10k) and filtrates were discarded. Classically, FASP would

Figure 1. Flowchart representing the different Amicon-adapted eFASP and FASP protocols compared in this study. A total of two FASP and eight Amicon-adapted eFASP protocols were compared on the basis of several performance characteristics. Apart from the purification protocol itself, protocols differed in regard to the type of filter unit used, the passivation of filter device, the enzymatic digestion and following cleanup. $n=3$ for FASP and Amicon-adapted eFASP without passivation; $n=2$ for Amicon-adapted eFASP with passivation.

have included an on-filter digestion. However, since our preliminary experiments using Amicons have shown suboptimal protein identifications and reproducibility with this approach, it was not further investigated and digestion has instead been performed in-solution. Precisely, concentrated samples were first recovered through a quick inverse spin of Amicons (1,000 $g/5$ min), thereafter transferred in a LoBind tube, diluted with 40 μ l of ABC, dosed for residual SDS, and trypsin-digested (ratio 1:100 w:w) overnight (37°C) in solution. Proteolysis was quenched with 2% final (v:v) FA. Peptide yields were assessed before and after solid phase extraction (SPE) with the BCA assay on a NanoDrop-2000c.

In parallel, 150- μ g aliquots of “non-SDS” lysate were also FASP-processed to confirm that nucleic acids and proteins do not interfere with the SDS colorimetric assay.

Amicon-adapted enhanced Filter-Aided Sample Preparation (eFASP)
(see [Supplementary Protocol](#))

When required by the protocol, filter units and collection tubes were filled with a 5% (v/v) Tween-20 solution in MS-grade water and soaked overnight at room temperature for passivation. Before use, passivated materials were twice immersed in water for 10 min with low-speed shaking. Filters and tubes were next twice filled with water and centrifuged at 14,000 g for 15 min (30k) or 25 min (10k).

A total of 20 fifty-microliter aliquots (150 μ g of proteins) of MEG-01 lysate were purified by the Amicon-adapted eFASP protocol, corresponding to 10 samples per type of filter (A10k and A30k) (see [Supplementary Table 3](#)). Lysates were mixed with iodoacetamide (50 mM final concentration), diluted to a volume of 500 μ l with 8 M urea in 0.1 M Tris-HCl pH 8 and kept in dark for a 30-min alkylation. Samples were then deposited onto filter devices for centrifugal concentration at 14,000 g for 15 min (30k) or 30 min (10k). The flow-through was discarded, as was systematically the case following each subsequent centrifugation. SDS removal was initiated by adding 500 μ l of 4% (w/v) NaDoc in 8 M urea to filter units, inverting devices thrice to mix the content well and centrifuging for 30 min

(30k) or 45 min (10k). Two more washing steps were performed with urea by filling devices to completeness, inverting filters and spinning for 15 min (30k) or 30 min (10k). Excess urea was then removed with two buffer exchanges during which devices were filled with ABC and spun for 20 min (30k) or 30 min (10k). Protein recovery prior to in-solution digestion was performed by an inverse spin of Amicons in new collection tubes (1,000 g/5 min), either passivated or not. Proteins were diluted to a final volume of 150 µl with ABC, mixed with trypsin/Lys-C Mix (ratio 1:25 w:w) in a LoBind tube and digested for 12 h at 37°C.

Samples were cleaned from residual SDS and NaDoc using either precipitation or phase transfer as described below. Yields were calculated with the BCA assay next to (1) precipitation or phase transfer and (2) SPE.

The concentration of SDS in purified samples was measured prior to digestion and next to precipitation or phase transfer. Before digestion, a first measurement of SDS was carried out after addition of 40 µl of ABC and a second after addition of the remaining ABC (150 µl final volume). Again, 150-µg aliquots of “non-SDS” control lysate were processed by Amicon-adapted eFASP to assess the stains-all colorimetric method’s specificity.

Precipitation and phase transfer

Following Amicon-adapted eFASP, two methods for simultaneous SDS and NaDoc removal were compared: precipitation and phase transfer. The precipitation protocol is based on Zhou *et al.*⁵ and Zhou *et al.*²⁴, with minor modifications. Briefly, samples were mixed with an equal volume of 4 M KCl, acidified with 2% FA and peptides were collected in the supernatant after centrifugation at 15,700 g for 15 min to pellet the potassium dodecyl sulfate (KDS) and deoxycholate precipitates.

The phase transfer procedure has been adapted from Masuda *et al.*²⁵ and Yeung and Stanley³¹. First, an equal volume of 4 M KCl was added to each sample, as well as 1 ml of water-saturated ethyl acetate. Samples were then centrifuged at 16,000 g for 1 min to allow separation of organic and aqueous phases. Part of the KDS-containing organic phase was discarded, leaving approximately 1 mm of the ethyl acetate layer to avoid loss of aqueous-soluble peptides. The extraction was repeated once with 600 µl of ethyl acetate before acidification with 2% FA, after which deoxycholic acid was removed by three rounds of ethyl acetate extraction. Samples were next vacuum dried for 10 min to remove the remaining organic solvent.

Residual sodium deoxycholate in Amicon-adapted eFASP-prepared samples

Unlike SDS, NaDoc cannot be quantified with the stains-all colorimetric assay. We therefore estimated its residual concentration in Amicon-adapted eFASP-prepared samples by visually comparing

their deoxycholate acid-precipitated pellet with various concentrations of acid-precipitated NaDoc (standard curve).

Processing limits of Amicon-adapted eFASP

To evaluate the processing capacities of Amicon-adapted eFASP regarding the highest amount of protein and the largest volume of sample that can be purified with no decrease in the effectiveness of SDS cleaning, we treated nine different samples containing various quantities of proteins and an equal number of volumes of lysate with A10k- and A30k-based Amicon-adapted eFASP. The remaining SDS concentration was quantified by the stains-all assay³⁰. ‘Protein’ experiments were carried out with the lowest volume of lysate possible (<60 µl) and amounts of proteins ranged between 100 and 600 µg as initially measured by BCA protein assay. For ‘volume’ experiments, 50 µg of protein extract solubilized in lysis buffer volumes varying between 30 and 300 µl were used.

Solid phase extraction

Recovered digested peptide mixtures were cleaned by SPE prior to LC-MS/MS analysis. SPE was carried out on an Oasis HLB 96-well µElution Plate containing 2 mg Sorbent (Waters, Milford, MA). Elution was performed with 75% acetonitrile containing 2% FA.

Liquid chromatography-tandem mass spectrometry (LC-MS/MS)

Acquisition was performed with a TripleTOF 5600 from AB SCIEX (Framingham, MA) equipped with an electrospray interface with a 25 µm I.D. capillary and coupled to an Eksigent µUHPLC (AB SCIEX). Analyst TF 1.6 software was used to control the instrument and for data processing and acquisition. The source voltage was set to 5.2 kV and temperature maintained at 375°C, curtain gas was set at 27 psi, gas one at 17 psi and gas two at 17 psi. Acquisition was performed in Data Dependent Acquisition (DDA) mode where a first period experiment was set for high resolution positive time-of-flight (TOF) MS in the m/z range 350–1,250 with a 25 ms accumulation time with declustering potential at 90 V and collision energy at 10. Top 40 ions were next triggered for 35 ms MS/MS experiment. Selected ions were next excluded after two occurrences for 45 sec. Separation was performed on a reversed phase HALO C18-ES column 0.3 µm i.d., 2.7 µm particles, 150 mm long (Advanced Materials Technology, Wilmington, DE) which was maintained at 50°C. Samples were injected by loop overfilling into a 5-µl loop that represented 3 µg of peptides. For the 60 min LC gradient, the mobile phase consisted in the following solvent A (0.2% v/v FA and 3% DMSO v/v in water) and solvent B (0.2% v/v FA and 3% DMSO in ethanol) at a flow rate of 5 µl/min. The gradient was the following: 0–45 min 2% B to 50% B, 45–49 min 50% B to 95% B, 49–53 min 95% B, 53–56 min 95% B to 2% B, 56–60 min 2% B and followed by a 2 min post-flush at 6 µl/min at final condition.

Data analysis

Protein identification was performed with ProteinPilot V4.5 beta (AB SCIEX) with the instrument pre-set for TripleTOF 5600, iodoacetamide as cysteine alkylation and urea denaturation as special factor. Thorough search with false discovery rate analysis was performed with biological modification emphasis against Human UniProtKB/Swiss-Prot Release 2014_01. For protein identification and data analysis using the validated Proteomics System Performance Evaluation Pipeline Software (PSPEP) from AB SCIEX, global false discovery rate was set at 1% and local false discovery rate was set at 5%. Description and review of this protein identification method can be found in Tang *et al.* (2008)³² and Serange and Noble (2012)³³. A combined ProteinPilot analysis of data acquired by each replicate of a specific protocol was also performed to generate a separate list of confidently identified proteins. Both combined and individual analyses of replicates are reported in the manuscript, the results of the latter being expressed unless otherwise indicated as: average $\pm$ standard deviation (SD) for the replicates. Note that in FASP and Amicon-adapted eFASP without passivation, data were summarized from triplicate experiments. In Amicon-adapted eFASP protocols with passivation, data were summarized from duplicate experiments. Chromatogram review and peak area information was performed with PeakView Software (AB SCIEX). Gene Ontology (GO) analyses were based on the PANTHER Classification System (v9.0) described by Mi *et al.*³⁴. A *Homo sapiens* genome-wide data analysis was performed on every protein ID lists generated by combined ProteinPilot analysis of replicates and accessions searched for 'Molecular Function', 'Biological Process' and 'Cellular Component' classes and subclasses.

We assessed the reproducibility of FASP and Amicon-adapted eFASP by manually inspecting the total ion currents (TIC) chromatograms of each protocol. For quantitative results, we extracted and cumulated, for each replicate, the intensity (area under the peak) of the five most intense peptides from the five most abundant identified proteins (25 peaks total) per protocol. Likewise, peptide retention time was compared between replicates. The identity of these five proteins was determined by the combined Protein Pilot analysis of each replicate's data. We then averaged the total cumulated intensity of the 25 peaks for every replicate, and compared protocols on the basis of the average cumulated peaks intensity and the resulting coefficient of variation (see [Equations 1](#)). We also performed the same analysis with the average peak height. Briefly, a cumulated peak height for each of the five proteins was calculated by adding up the peak height of each of its five most intense peptides. For every single replicate, a global peak height was obtained by averaging the cumulated peak height of the five proteins. Then, the global peak heights were averaged for all protocols (see [Equations 2](#)).

Equations 1: Average cumulated signals intensity. The five most intense proteins along with their five most intense peptides for a

given protocol were selected according to the combined Protein Pilot analysis of each of its replicate's data.

(1) Protein signal intensity = intensity of peptide 1 + intensity of peptide 2 + intensity of peptide 3 + intensity of peptide 4 + intensity of peptide 5

(2) Replicate cumulated signal intensity = Protein signal intensity 1 + Protein signal intensity 2 + Protein signal intensity 3 + Protein signal intensity 4 + Protein signal intensity 5

(3) Protocol average cumulated signals intensity = (cumulated signal intensity of R1 + cumulated signal intensity of R2 + cumulated signal intensity of R3)/3

Equations 2: Average peak height. The five most intense proteins along with their five most intense peptides for a given protocol were selected according to the combined Protein Pilot analysis of each of its replicate's data.

(1) Protein cumulated peak height = peak height of peptide 1 + peak height of peptide 2 + peak height of peptide 3 + peak height of peptide 4 + peak height of peptide 5

(2) Replicate global peak height = (Cumulated peak height of protein 1 + protein 2 + protein 3 + protein 4 + protein 5)/5

(3) Protocol global peak height = (global peak height of R1 + global peak height of R2 + global peak height of R3)/3

Results and discussion

Few reliable and reproducible sample preparation methods are available for proteomics-based biomarker discovery, especially when conducted with SDS-solubilized biological samples. FASP is certainly one of the most efficient and simple available methods for such experiments, especially as it has been successfully applied to a variety of different samples including clinical material¹⁰⁻¹². Nevertheless, as is the case with every other sample preparation method, FASP is not a processing method that works optimally with all biological samples and in every experimental setting¹⁵⁻²⁰. In light of the very extensive literature on 'best practices' in sample preparation for proteomics experiments, we sought to tailor the FASP workflow for the Amicon filter device in order to improve its analytical performances. In this regard, we designed the Amicon-adapted eFASP, a complementary sample preparation method that introduces key modifications to the FASP and that is intended to improve proteome coverage and reproducibility. Among these modifications are: a passivation of the filter unit for improving SDS removal and peptide yields, a critical cleaning step with a high-content NaDoc buffer, and the substitution of on-filter for in-solution 'one-step-two-enzymes' trypsin/Lys-C digestion^{35,36}. In this study, we extensively characterized the analytical properties of the Amicon-adapted eFASP by comparing it to the classic one-step digestion FASP approach⁹ during shotgun proteomics experiments.

The extensive characterization of the human cancer cell MEG-01 proteome (a model of human leukemic megakaryoblast) is also described for the first time herein.

Effect of SDS on digestion efficacy

In agreement with Zhou *et al.*⁵, we confirmed the extreme sensitivity of trypsin to low concentrations of SDS, with complete digestion being achieved for samples containing $\leq 0.05\%$ SDS (Figure 2A and Supplementary Figure 1A). The upper tolerable SDS limit is however not known for combined trypsin and Lys-C. Specifically, we found that optimal trypsin/Lys-C digestion efficacy, in a ratio of 1:25 w/w as is used in the Amicon-adapted eFASP protocol, was achieved

for SDS concentrations $\leq 0.07\%$ and that this combined proteolysis performed better than trypsin alone in higher SDS concentrations (Figure 2B and Supplementary Figure 1B), likely due to the greater stability of Lys-C in denaturing conditions²⁵. The enzyme-to-protein ratio of 1:25 w/w in Amicon-adapted eFASP, representing a four-fold increase as compared to FASP, was chosen because it was previously shown preferable in comprehensive proteomics¹⁶.

Removal of SDS

SDS is not easily removed from proteins to which it binds strongly through ionic and hydrophobic interactions⁵. In FASP, usage of a washing buffer composed solely of urea has been assumed to be

Figure 2. Comparison of residual SDS concentrations and peptide yields in Amicon-adapted eFASP and FASP protocols. Effect of increasing SDS concentrations on the completion of overnight (A) trypsin (ratio 1:100 w/w) and (B) trypsin/Lys-C (ratio 1:25 w/w) digestion of MEG-01 total cell lysates assessed by a Coomassie-stained 12% SDS-PAGE. (C) Concentrations of residual SDS prior to digestion and (D) peptide yields in FASP (n=3) and Amicon-adapted eFASP (n=10). Nonparametric Mann-Whitney U test was performed for comparison of mean SDS concentrations and peptide yields: ** $p < 0.01$, * $p < 0.05$. Mean ($\pm$ S.D.) Statistical comparisons were made between the same filters as used in FASP and Amicon-adapted eFASP. TCL: Total cell lysate.

sufficient to completely deplete the SDS from up to 30 μ l of a small protein extract (≤ 250 μ g)^{9,13,14}. However, this assumption was supported by the measurement of the residual detergent concentration in a FASP-cleaned 100- μ l 2% SDS protein-free lysis buffer¹⁴. Rationally, the processing of a cell lysate would have been more appropriate for this purpose, especially as proteins are strongly bonded to SDS and may lower its critical micelle concentration (CMC ~ 6 –8 mM)³⁷, which can greatly impede its elimination (<http://www.millipore.com/techpublications/tech1/6djp7f>, EMD Millipore). Indeed, SDS micelles are not readily filtered by 30k and 10k MWCO membranes as their molecular weight is ~ 18 –40 kDa¹⁴. Furthermore, urea is an uncharged chaotrope that disrupts hydrogen bonds and hydrophobic molecular interactions³⁸. It enables the dissociation of SDS nonpolar tails from proteins but may not be as effective for disrupting the ionically bonded sulfate heads. We reasoned that introduction of the weak anionic surfactant NaDoc in the Amicon-adapted eFASP workflow would enhance the disruption of these powerful ionic interactions. Therefore, we measured for the first time the residual SDS concentration in Amicon-adapted eFASP- and FASP-processed 150- μ g protein extracts. Owing to incorporation of a cleaning step with a high-concentration 4% (w/v) NaDoc buffer, Amicon-adapted eFASP significantly improves the removal of SDS from high-content protein samples independently of the filter used, and diluting proteins in a final volume of digestion buffer of 150 μ l was sufficient to lower SDS concentration below the trypsin-activity-compatible 0.05% threshold for all conditions (Figure 2C and Supplementary Table 4). We also found that 4% NaDoc as used in the present Amicon-adapted eFASP version was more effective than 0.2% deoxycholic acid as is used in Erde's eFASP¹⁸ for removing the SDS (Supplementary Table 5). In fact, among the ten different NaDoc concentrations (ranging from 0.1 to 5%) tested during Amicon-adapted eFASP designing, a single wash with 4% NaDoc was the most effective approach to deplete the SDS. Moreover, as exemplified by the complete depletion of SDS from FASP-processed samples containing <100 μ g of proteins, the efficiency of filter-based SDS removal is tightly related to the initial amount of protein in the sample. This is further supported by the fact that recent experiments in which FASP led to lower-than-expected protein identifications have used >100 μ g of proteins^{15–19}. Equally important for efficient depletion of SDS is the starting volume of the sample, which was different in Amicon-adapted eFASP (50 μ l) and FASP (30 μ l). In fact, samples must be sufficiently diluted during buffer exchanges to minimize formation of unfilterable SDS micelles. This is more readily done with Amicon-adapted eFASP as the buffer volume is beneficially increased to 500 μ l throughout the protocol, relative to the 200 μ l in the FASP. Thereupon, when purifying <30 μ l of a 150- μ g protein extract with Amicon-adapted eFASP, SDS is almost completely

($<0.05\%$) removed with both 10k and 30k filters. Of note, modifications brought by the Amicon-adapted eFASP protocol extend the processing limits to 50 μ l/400 μ g and 125 μ l/600 μ g for 10k and 30k filters, respectively (volume/protein) (Supplementary Table 6, Supplementary Table 7).

Minimizing nonspecific protein and peptide binding to filter device's plastic walls and membrane by passivation is possible with Tween-20 and SDS¹⁸. This reasonably implies that SDS present in the sample may bind to non-passivated filter reservoir's surfaces and may not be completely removed despite buffer exchanges, further explaining the difficulty of removing it from spin filter-processed proteins. This may however be prevented by pre-treating filters with Tween-20. As such, lower concentrations of SDS were achieved in samples eFASP-processed with passivated Amicon filters (Supplementary Table 4). In sum, Amicon-adapted eFASP efficiently eliminates SDS by combining a passivation of the filter device, an increase of exchange buffer volume and an incorporation of a high-concentration NaDoc-based cleaning step. These features respectively prevent non-specific binding of SDS to device's surfaces, minimize formation of unfilterable SDS micelles and ease dissociation of ionically bonded SDS from basic amino acids.

Peptide yields

Amicon-adapted eFASP provided 10% higher peptide recoveries for each type of filter as compared to FASP (Figure 2D and Supplementary Figure 2). The better yields of Amicon-adapted eFASP may be attributed to the filter passivation (Supplementary Table 4) and to its reduced number of centrifugations.

Proteome coverage

We next aimed to evaluate if Amicon-adapted eFASP compares favourably to FASP in regard to depth of proteome coverage. We found that Amicon-adapted eFASP noticeably deepened the MEG-01 cell's proteome coverage relative to FASP (Figure 3A, B and Table 1). The lower concentrations of SDS remaining in samples after Amicon-adapted eFASP protocols, which were lower than the trypsin/Lys-C activity-compatible threshold, may in part explain the difference in proteome coverage. A considerable number of peptides/proteins were confidently identified (1% false discovery rate) in the three replicates of the Amicon-adapted eFASP/A10k/PPT protocol: 6,479/1,595, 6,567/1,596 and 6,424/1,603. The delivery rate was accordingly very high, with 27 proteins identified per minute, and the combined peptides-to-protein ratio was close to 6.0, an increment of 1 peptide per protein relative to FASP. This Amicon-adapted eFASP protocol (i.e. A10k/PPT) is the least variable regarding the number of identified peptides (coefficient of variation [CV] = 1.1%) and proteins (CV = 0.28%)

Figure 3. Comprehensive proteomics analysis of human leukemic megakaryoblast (MEG-01) lysates by Amicon-adapted eFASP and FASP approaches. (A) Peptide and (B) protein identifications achieved by Amicon-adapted eFASP/precipitation and FASP protocols at a 1% global false discovery rate. Displayed are the results of the combined (green) and individual (blue) analyses of replicates. For individual analyses, data corresponding to the mean $\pm$ standard deviation were summarized from triplicate experiments. Comparison of peptide and protein total identifications between Amicon-adapted eFASP and FASP was performed with the nonparametric Mann-Whitney U test. (C) Protein identifications overlap from combined analysis of Amicon-adapted eFASP/precipitation and FASP protocols. The total number of proteins identified per protocol is indicated in the corresponding box. (D) Comparison of Gene Ontology annotations for specific cellular component classes in Amicon-adapted eFASP/precipitation- and FASP-processed and analyzed SDS lysates. Bars represent the percentage of identified proteins with the indicated annotations. PPT: Precipitation.

Table 1. LC-MS/MS analysis of MEG-01 total cell lysates prepared by the different FASP and Amicon-adapted eFASP protocols.

	Protocol	Peptide Identifications		Protein Identifications		Average Sequence Coverage (%)		Molecular Weight (kDa)		Peptides-to-protein Ratio	
		Total ^a	Average (±SD) ^b	Total ^a	Average (±SD) ^b	Combined analysis ^a	Individual analysis (±SD) ^b	Median ^a	Average (±SD) ^b	Combined analysis ^a	Average ^b
FASP	Amicon 10k	5,755	3,080 (174.7)	1,199	851 (57.2)	23.96	18.83 (0.22)	43.90	62.43 (127.70)	4.8	3.6
	Amicon 30k	6,223	3,299 (802.9)	1,457	1,010 (102.8)	20.38	15.27 (1.87)	46.99	61.62 (57.23)	4.3	3.2
Amicon-based eFASP	Phase	9,907	5,153 (354.9)	1,683	1,214 (87.1)	25.84	20.64 (1.29)	44.76	58.61 (54.32)	5.9	4.2
	PPT	12,431	6,490 (72.1)	2,177	1,598 (4.4)	22.22	18.29 (0.24)	46.42	61.24 (56.49)	5.7	4.1
	PT20/Phase	8,171	5,262 (145.0)	1,525	1,208 (31.8)	24.15	19.74 (0.79)	46.39	60.63 (57.13)	5.4	4.4
	PT20/PPT	10,338	6,716 (5.0)	2,046	1,680 (25.5)	19.68	17.58 (0.21)	46.90	64.02 (105.60)	5.1	4.0
	Phase	10,343	4,313 (1,569.1)	1,677	1,075 (261.5)	25.01	16.94 (2.78)	45.49	60.30 (59.72)	6.2	3.9
	PPT	11,167	6,007 (385.3)	1,895	1,406 (80.0)	23.10	18.47 (0.19)	46.41	59.89 (55.93)	5.9	4.3
Amicon 30k	PT20/Phase	7,134	4,643 (270.1)	1,437	1,131 (97.6)	21.39	17.83 (0.57)	46.87	60.41 (54.20)	5.0	4.1
	PT20/PPT	9,805	6,395 (35.4)	1,826	1,508 (5.0)	20.71	18.25 (0.04)	46.42	59.88 (54.95)	5.4	4.2

^a. Results of the combined analysis of replicates.^b. Average ± standard deviation of data acquired by individual analysis of replicates.

n=3 for FASP and Amicon-adapted eFASP without passivation; n=2 for Amicon-adapted eFASP with passivation.

Phase, Phase transfer; PPT, Precipitation; PT20, 5% Tween-20 passivation.

per replicate. Combined analysis led to the identification of 12,431 peptides and 2,177 proteins, a significant increase of 200% for peptides and 150% for proteins compared to FASP with a 30k MWCO filter. Moreover, the protein overlap is $\geq 70\%$ between the Amicon-adapted eFASP/A10k/PPT protocol and every other single one (Figure 3C and Supplementary Table 8). Along with improvements in SDS removal and peptide recoveries, passivation increased peptide and proteins identifications by 7% and 6%, respectively. There is also $\geq 80\%$ homology among the proteins identified by a passivated Amicon-adapted eFASP approach and its non-passivated equivalent. In conjunction with the Amicon-adapted eFASP/A10k/PPT protocol, passivation yields identification of 1,662 and 1,698 proteins, for the two replicates. In light of the previous results, it seems clear that passivation of filter, simultaneous trypsin/Lys-C in-solution digestion and incorporation of NaDoc to enhance both removal of SDS and enzyme activity, were beneficent for the workflow. In fact, the concentration of residual NaDoc in Amicon-adapted eFASP digestion buffer as assessed by precipitation ranged between 0.20% and 0.35%. This concentration is known to significantly improve digestion reproducibility and to facilitate proteolysis of membrane and hydrophobic proteins^{22,25,39}. As logically expected with a lower MWCO filter, fewer proteins were lost in A10k-based Amicon-adapted eFASP protocols, enabling the identification of more proteins. The opposite finding in FASP might be explained by the lower digestion efficiency achieved for A10k-processed samples due to higher concentrations of residual SDS, leading to loss of incompletely cleaved polypeptides during SPE.

Finally, it is worth noting that phase transfer-based Amicon-adapted eFASP protocols yielded 20% lower protein identifications as compared to precipitation-based ones (Table 1). This is in good agreement with Lin *et al.*³⁹ and might have been caused by loss of ethyl acetate-miscible hydrophobic peptides and incorporation of peptide-containing water droplets in the organic phase during the extractions³¹. Moreover, there is no significant difference in sequence coverage and mean molecular weight of identified proteins among protocols (Supplementary Figure 3–Supplementary Figure 5 and Supplementary Table 4).

Amicon-adapted eFASP is the first method to enable the identification in only one hr of nearly 1,700 proteins in a single analysis and 2,200 proteins in three separate runs from minimally prepared and non-fractionated SDS-solubilized lysates. To our knowledge, this is also the first study to achieve an identification yield this high in one hr with a 15-cm column, favourably comparing to Hebert *et al.*⁴⁰ that reached 3,977 identifications with a 35-cm column.

The complete list of peptides and protein IDs identified by each protocol with both combined and individual analyses of replicates can be found in Supplementary Dataset 1–Supplementary Dataset 11.

Dataset 1. Supplementary Datasets 1–11

<http://dx.doi.org/10.5256/f1000research.6529.d102594>

- Supplementary Dataset 1: Proteins identified by combined analysis of replicates for each of the ten protocols of the study
- Supplementary Dataset 2: Peptides and proteins identified by individual analysis of each replicate for the FASP, Amicon 10k protocol
- Supplementary Dataset 3: Peptides and proteins identified by individual analysis of each replicate for the FASP, Amicon 30k protocol
- Supplementary Dataset 4: Peptides and proteins identified by individual analysis of each replicate for the Amicon-adapted eFASP, Amicon 10k, phase transfer protocol
- Supplementary Dataset 5: Peptides and proteins identified by individual analysis of each replicate for the Amicon-adapted eFASP, Amicon 10k, precipitation protocol
- Supplementary Dataset 6: Peptides and proteins identified by individual analysis of each replicate for the Amicon-adapted eFASP, passivated Amicon 10k, phase transfer protocol
- Supplementary Dataset 7: Peptides and proteins identified by individual analysis of each replicate for the Amicon-adapted eFASP, passivated Amicon 10k, precipitation protocol
- Supplementary Dataset 8: Peptides and proteins identified by individual analysis of each replicate for the Amicon-adapted eFASP, Amicon 30k, phase transfer protocol
- Supplementary Dataset 9: Peptides and proteins identified by individual analysis of each replicate for the Amicon-adapted eFASP, Amicon 30k, precipitation protocol
- Supplementary Dataset 10: Peptides and proteins identified by individual analysis of each replicate for the Amicon-adapted eFASP, passivated Amicon 30k, phase transfer protocol
- Supplementary Dataset 11: Peptides and proteins identified by individual analysis of each replicate for the Amicon-adapted eFASP, passivated Amicon 30k, precipitation protocol⁴¹

Methods' reproducibility

Visual appraisal of TIC identified Amicon-adapted eFASP/A10k/PPT as the most reproducible protocol, which is further improved for significant signal enhancement throughout the gradient by passivation (Figure 4, Figure 5 and Supplementary Figure 6D). Noteworthy, the TIC spectra of each replicate are almost perfectly superposed one to another in these two protocols as exemplified by the $\sim 6.5\%$ CV in the cumulated intensity and global peak height. In fact, combination of precipitation and passivation always provided the strongest signal and the highest reproducibility for both Amicon 10k and 30k when used in the Amicon-adapted eFASP workflow (Supplementary Figure 6–Supplementary Figure 8). Signal suppression from remaining SDS seems to be present in both FASP/A10k and FASP/A30k chromatograms, which might also have contributed to the lower

Figure 4. Qualitative assessment of reproducibility of Amicon-adapted eFASP Amicon 10k precipitation and FASP Amicon 10k protocols. Total ion current (TIC) chromatograms of the three replicates from **(A)** Amicon-adapted eFASP Amicon 10k precipitation (maximal signal intensity: 4.0e7) and **(C)** FASP Amicon 10k (maximal signal intensity: 1.5e7) within the total 1-hr analytical liquid chromatography gradient. Extracted-ion chromatograms of Histone H2B peptide LLLPGELAK from **(B)** Amicon-adapted eFASP Amicon 10k precipitation (maximum signal intensity: 2.5e6) and **(D)** FASP Amicon 10k (maximum signal intensity: 7.5e5). **(B, D)** The retention time (minutes) of the peptide is also indicated for each replicate.

identification yields of FASP (Figure 4C and Supplementary Figure 6A). Indeed, it was previously shown that $\geq 0.01\%$ (w/v) SDS significantly suppresses analyte ion signals and dramatically reduces protein identification in bottom-up proteomics approaches⁴.

Gene Ontology

Broad and unbiased proteomics representation of subcellular compartments is a prerequisite for successful protein biomarker discovery. There were some differences in Gene Ontology (GO) cell component annotations between FASP and Amicon-adapted eFASP protocols, with more membrane and mitochondrion proteins, as well as more protein complex annotated in the latter (Figure 3D). On the other hand, the Amicon-adapted eFASP/passivated A10k/PPT protocol provided the highest combined GO annotations for nuclear and membrane proteins (Supplementary Figure 9), likely owing to

the presence of NaDoc in the digestion buffer. Indeed, NaDoc helps to solubilize and expose membrane proteins' cleavage sites to trypsin²⁶, leading to an increase of their identification as previously shown^{18,23–25,39}.

In summary, our Amicon-adapted eFASP protocol is an optimized Amicon unit-based sample preparation approach that provides efficient removal of SDS, high reproducibility, deep proteome coverage and enhanced identification of membrane proteins. Amicon-adapted eFASP enabled the identification of 1,700 proteins from SDS-solubilized proteins with minimal sample preparation and no fractionation, within a single run one-hr liquid chromatography and using a 15-cm column. Its implementation in proteomics may therefore improve the identification of high-confidence and significant biomarkers.

Figure 5. Quantitative assessment of reproducibility of selected Amicon-adapted eFASP and FASP approaches. Bars show the average cumulated MS signals intensities from each replicate for the five most intense peptides from the five most abundant identified proteins per protocol. Error bars are representative of the standard deviation of the mean associated with each replicate analysis. Data were summarized from triplicate experiments and duplicate experiments for the Amicon-adapted eFASP passivated Amicon 10k and Amicon 30k precipitation (Amicon-adapted eFASP/A10k/PT20/PPT and Amicon-adapted eFASP/A30k/PT20/PPT) protocols. Numbers above error bars indicate the coefficient of variation in the average cumulated MS signal intensities. PPT: Precipitation. PT20: 5% Tween-20 passivation.

Data availability

F1000Research: Dataset 1. [Supplementary Dataset 1–Supplementary Dataset 11](#), [10.5256/f1000research.6529.d102594](#)⁴¹

Author contributions

DP designed and developed the method, designed and interpreted the experiments, performed the experiments and the bioinformatics analyses, and wrote the manuscript. HG contributed in the development of the method, designed and interpreted the experiments, performed the mass spectrometry experiments, assisted in data analysis and reviewed the manuscript. JD and FC supervised the work, designed and interpreted the experiments, assisted in data analysis and reviewed the manuscript.

Competing interests

No competing interests were disclosed.

Grant information

This work was supported by Université de Sherbrooke and the CHUS Research Center. DP holds a Frederick Banting and Charles Best Canada Graduate Scholarship Master Award from the Canadian Institutes of Health Research (CIHR) and an M.D.-M.Sc. Graduate Scholarship Master Award from the *Fonds de Recherche du Québec – Santé* (FRQS). HG holds a MITACS accelerated postdoctoral scholarship.

I confirm that the funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

We gratefully thank Dr. Robert Day for his participation in the elaboration and critical review of the manuscript. We also thank Sandra Gagnon for technical support, Simon Pellerin for data processing, and Artuela Çaku and Paméla Bouvier for helpful comments and suggestions on this manuscript.

Supplementary material

Supplementary Material Available:

- Supplementary Figure 1. Effect of increasing concentrations of SDS on FASP (trypsin) and Amicon-adapted eFASP (trypsin/Lys-C) overnight digestion efficacy
- Supplementary Figure 2. Peptide yields
- Supplementary Figure 3. Sequence coverage by intervals
- Supplementary Figure 4. Miscleavage frequencies
- Supplementary Figure 5. Frequency of identified proteins' theoretical molecular weight (kDa) for each protocol investigated
- Supplementary Figure 6. Total ion current chromatograms
- Supplementary Figure 7. Quantitative assessment of reproducibility: average cumulated signals intensity
- Supplementary Figure 8. Quantitative assessment of reproducibility: average peaks height
- Supplementary Figure 9. Distribution of Gene Ontology (GO) annotations of human leukemic megakaryoblast (MEG-01) proteins identified by Amicon-adapted eFASP and FASP protocols (combined analysis of replicates)
- Supplementary Table 1. Spectrophotometric quantification of SDS with the stains-all colorimetric assay
- Supplementary Table 2. Minimal effect of solutions and reagents used in FASP and Amicon-adapted eFASP protocols on stains-all dosage solution absorbance at 438 nm
- Supplementary Table 3. Summary of the 10 protocols compared in this study
- Supplementary Table 4. LC-MS/MS results of FASP and Amicon-adapted eFASP protocols
- Supplementary Table 5. Concentration of residual SDS in 0.2% deoxycholic acid-based enhanced FASP (eFASP)-processed MEG-01 whole cell lysates
- Supplementary Table 6. Volume processing capacity of Amicon-adapted eFASP
- Supplementary Table 7. Protein processing capacity of Amicon-adapted eFASP
- Supplementary Table 8. Protein identifications overlap
- Supplementary Protocol. Amicon-adapted enhanced Filter-Aided Sample Preparation (eFASP)

[Click here to access the data.](#)

References

1. Anderson NL: **The clinical plasma proteome: a survey of clinical assays for proteins in plasma and serum.** *Clin Chem.* 2010; **56**(2): 177–185.
[PubMed Abstract](#) | [Publisher Full Text](#)
2. Capelo JL, Carreira R, Diniz M, *et al.*: **Overview on modern approaches to speed up protein identification workflows relying on enzymatic cleavage and mass spectrometry-based techniques.** *Anal Chim Acta.* 2009; **650**(2): 151–159.
[PubMed Abstract](#) | [Publisher Full Text](#)
3. Bodzon-Kulakowska A, Bierczynska-Krzysik A, Dylag T, *et al.*: **Methods for samples preparation in proteomic research.** *J Chromatogr B Analyt Technol Biomed Life Sci.* 2007; **849**(1–2): 1–31.
[PubMed Abstract](#) | [Publisher Full Text](#)
4. Botelho D, Wall MJ, Vieira DB, *et al.*: **Top-down and bottom-up proteomics of SDS-containing solutions following mass-based separation.** *J Proteome Res.* 2010; **9**(6): 2863–2870.
[PubMed Abstract](#) | [Publisher Full Text](#)
5. Zhou JY, Dann GP, Shi T, *et al.*: **Simple sodium dodecyl sulfate-assisted sample preparation method for LC-MS-based proteomics applications.** *Anal Chem.* 2012; **84**(6): 2862–2867.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. Rundlett KL, Armstrong DW: **Mechanism of signal suppression by anionic surfactants in capillary electrophoresis-electrospray ionization mass spectrometry.** *Anal Chem.* 1996; **68**(19): 3493–3497.
[PubMed Abstract](#) | [Publisher Full Text](#)
7. Hustoft HK, Malerod H, Wilson SR, *et al.*: **A Critical Review of Trypsin Digestion for LC-MS Based Proteomics.** In: Man T.K. & Flores, R.J., eds. *Integrative Proteomics.* Chapter 4. InTECH, Rijeka, Croatia. 2012.
[Publisher Full Text](#)
8. Manza LL, Stamer SL, Ham AJ, *et al.*: **Sample preparation and digestion for proteomic analyses using spin filters.** *Proteomics.* 2005; **5**(7): 1742–1745.
[PubMed Abstract](#) | [Publisher Full Text](#)
9. Wiśniewski JR, Zougman A, Nagaraj N, *et al.*: **Universal sample preparation method for proteome analysis.** *Nat Methods.* 2009; **6**(5): 359–362.
[PubMed Abstract](#) | [Publisher Full Text](#)
10. Wiśniewski JR, Ostasiewicz P, Mann M: **High recovery FASP applied to the proteomic analysis of microdissected formalin fixed paraffin embedded cancer tissues retrieves known colon cancer markers.** *J Proteome Res.* 2011; **10**(7): 3040–3049.
[PubMed Abstract](#) | [Publisher Full Text](#)

11. Wiśniewski JR, Nagaraj N, Zougman A, *et al.*: **Brain phosphoproteome obtained by a FASP-based method reveals plasma membrane protein topology.** *J Proteome Res.* 2010; 9(6): 3280–3289.
[PubMed Abstract](#) | [Publisher Full Text](#)
12. Geiger T, Cox J, Ostasiewicz P, *et al.*: **Super-SILAC mix for quantitative proteomics of human tumor tissue.** *Nat Methods.* 2010; 7(5): 383–385.
[PubMed Abstract](#) | [Publisher Full Text](#)
13. Liebler DC, Ham AJ: **Spin filter-based sample preparation for shotgun proteomics.** *Nat Methods.* 2009; 6(11): 785–786.
[PubMed Abstract](#) | [Publisher Full Text](#)
14. Wiśniewski JR, Zielinska DF, Mann M: **Comparison of ultrafiltration units for proteomic and N-glycoproteomic analysis by the filter-aided sample preparation method.** *Anal Biochem.* 2011; 410(2): 307–309.
[PubMed Abstract](#) | [Publisher Full Text](#)
15. León IR, Schwämmle V, Jensen ON, *et al.*: **Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis.** *Mol Cell Proteomics.* 2013; 12(10): 2992–3005.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
16. Hustoft HK, Reubsaet L, Greibrokk T, *et al.*: **Critical assessment of accelerating trypsination methods.** *J Pharm Biomed Anal.* 2011; 56(5): 1069–1078.
[PubMed Abstract](#) | [Publisher Full Text](#)
17. Bereman MS, Egerton JD, MacCoss MJ: **Comparison between procedures using SDS for shotgun proteomic analyses of complex samples.** *Proteomics.* 2011; 11(14): 2931–2935.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Erde J, Loo RR, Loo JA: **Enhanced FASP (eFASP) to increase proteome coverage and sample recovery for quantitative proteomic experiments.** *J Proteome Res.* 2014; 13(4): 1885–1895.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Vowinckel J, Capuano F, Campbell K, *et al.*: **The beauty of being (label)-free: sample preparation methods for SWATH-MS and next-generation targeted proteomics.** *F1000Res.* 2013; 2: 272.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Atrih A, Mudaliar MA, Zakikhani P, *et al.*: **Quantitative proteomics in resected renal cancer tissue for biomarker discovery and profiling.** *Br J Cancer.* 2014; 110(6): 1622–1633.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
21. Liebler DC, Ham AJ: **Spin filter-based sample preparation for shotgun proteomics.** *Nat Methods.* 2009; 6(11): author reply 785–786.
[PubMed Abstract](#) | [Publisher Full Text](#)
22. Proc JL, Kuzyk MA, Hardie DB, *et al.*: **A quantitative study of the effects of chaotropic agents, surfactants, and solvents on the digestion efficiency of human plasma proteins by trypsin.** *J Proteome Res.* 2010; 9(10): 5422–5437.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. Lin Y, Zhou J, Bi D, *et al.*: **Sodium-deoxycholate-assisted tryptic digestion and identification of proteolytically resistant proteins.** *Anal Biochem.* 2008; 377(2): 259–266.
[PubMed Abstract](#) | [Publisher Full Text](#)
24. Zhou J, Zhou T, Cao R, *et al.*: **Evaluation of the application of sodium deoxycholate to proteomic analysis of rat hippocampal plasma membrane.** *J Proteome Res.* 2006; 5(10): 2547–2553.
[PubMed Abstract](#) | [Publisher Full Text](#)
25. Masuda T, Tomita M, Ishihama Y: **Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis.** *J Proteome Res.* 2008; 7(2): 731–740.
[PubMed Abstract](#) | [Publisher Full Text](#)
26. Chen EI, Cociorva D, Norris JL, *et al.*: **Optimization of mass spectrometry-compatible surfactants for shotgun proteomics.** *J Proteome Res.* 2007; 6(7): 2529–2538.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Strader MB, Tabb DL, Hervey WJ, *et al.*: **Efficient and specific trypsin digestion of microgram to nanogram quantities of proteins in organic-aqueous solvent systems.** *Anal Chem.* 2006; 78(1): 125–134.
[PubMed Abstract](#) | [Publisher Full Text](#)
28. Klammer AA, MacCoss MJ: **Effects of modified digestion schemes on the identification of proteins from complex mixtures.** *J Proteome Res.* 2006; 5(3): 695–700.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Nel AJ, Garnett S, Blackburn JM, *et al.*: **Comparative reevaluation of FASP and enhanced FASP methods by LC-MS/MS.** *J Proteome Res.* 2015; 14(3): 1637–1642.
[PubMed Abstract](#) | [Publisher Full Text](#)
30. Rusconi F, Valton E, Nguyen R, *et al.*: **Quantification of sodium dodecyl sulfate in microliter-volume biochemical samples by visible light spectroscopy.** *Anal Biochem.* 2001; 295(1): 31–37.
[PubMed Abstract](#) | [Publisher Full Text](#)
31. Yeung YG, Stanley ER: **Rapid detergent removal from peptide samples with ethyl acetate for mass spectrometry analysis.** *Curr Protoc Protein Sci.* 2010; Chapter 16: Unit 16.12.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. Tang WH, Shilov IV, Seymour SL: **Nonlinear fitting method for determining local false discovery rates from decoy database searches.** *J Proteome Res.* 2008; 7(9): 3661–3667.
[PubMed Abstract](#) | [Publisher Full Text](#)
33. Serang O, Noble W: **A review of statistical methods for protein identification using tandem mass spectrometry.** *Stat Interface* 2012; 5(1): 3–20.
[PubMed Abstract](#) | [Free Full Text](#)
34. Mi H, Muruganujan A, Casagrande JT, *et al.*: **Large-scale gene function analysis with the PANTHER classification system.** *Nat Protoc.* 2013; 8(8): 1551–1566.
[PubMed Abstract](#) | [Publisher Full Text](#)
35. Saveliev S, Bratz M, Zubarev R, *et al.*: **Trypsin/Lys-C protease mix for enhances protein mass spectrometry analysis.** *Nat Methods.* 2013.
[Reference Source](#)
36. Glatter T, Ludwig C, Ahmé E, *et al.*: **Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion.** *J Proteome Res.* 2012; 11(11): 5145–5156.
[PubMed Abstract](#) | [Publisher Full Text](#)
37. Chattopadhyay A, London E: **Fluorimetric determination of critical micelle concentration avoiding interference from detergent charge.** *Anal Biochem.* 1984; 139(2): 408–412.
[PubMed Abstract](#) | [Publisher Full Text](#)
38. Cañas B, Piñeiro C, Calvo E, *et al.*: **Trends in sample preparation for classical and second generation proteomics.** *J Chromatogr A.* 2007; 1153(1–2): 235–258.
[PubMed Abstract](#) | [Publisher Full Text](#)
39. Lin Y, Liu Y, Li J, *et al.*: **Evaluation and optimization of removal of an acid-insoluble surfactant for shotgun analysis of membrane proteome.** *Electrophoresis.* 2010; 31(16): 2705–2713.
[PubMed Abstract](#) | [Publisher Full Text](#)
40. Hebert AS, Richards AL, Bailey DJ, *et al.*: **The one hour yeast proteome.** *Mol Cell Proteomics.* 2014; 13(1): 339–347.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
41. Pellerin D, Gagnon H, Dubé J, *et al.*: **Dataset 1 in: Amicon-adapted enhanced FASP: an in-solution digestion-based alternative sample preparation method to FASP.** *F1000Research.* 2015.
[Data Source](#)

Open Peer Review

Current Peer Review Status: ? ?

Version 2

Reviewer Report 19 February 2016

<https://doi.org/10.5256/f1000research.7632.r8891>

© 2016 Fricker L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lloyd D. Fricker

Department of Molecular Pharmacology, Albert Einstein College of Medicine of Yeshiva University, Bronx, NY, USA

The authors made some superficial changes to version 1, but version 2 still have problems with the supplemental data table. There are columns with no values! And they still list proteins identified with "0" peptides. If no peptides were confidently identified, how can the protein be identified? Before drawing conclusions from the manuscript, readers should consult the supplementary data files and make their own conclusions.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 01 February 2016

<https://doi.org/10.5256/f1000research.7632.r11760>

© 2016 Bonneil É. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Éric Bonneil

Institute for Research in Immunology and Cancer, University of Montreal, Montréal, QC, Canada

Amicon-adapted enhanced FASP: an in-solution digestion-based alternative sample preparation method to FASP.

The topic of SDS removal is, to my opinion, of a great interest since SDS is essential to a lot of proteomic sample preparation. And it happens more often than not that SDS could not be avoided. The only other option than FASP is to run a SDS-PAGE gel, to do a protein precipitation. Those avenues have shown some limitations in terms of sample recovery. So there is some space for improvements in SDS-removal sample preparation for proteomics. In this paper, the authors have adapted the method developed by Erde J. et al to Amicon filter units. While not dramatically new, this could be helpful for scientists willing to optimize their proteome recovery from SDS-containing samples.

The experimental part is carefully written with a lot of checkpoints. The authors clearly changed a lot of different parameters and came with reasonable explanations for the results. One small comment, the HALO C18 column could not be 0.3 μ m id.

I am a bit puzzled by the lack of overlap between all the 4 experiments presented in Figure 3c. For example, we could think that the use of eFASP A10k will yield less proteins because the SDS removal is less effective than with eFASP30K. Indeed, even if the SDS level is still below the critical level for the Lys-C/ trypsin digestion, it might affect the performance of LC-MS. But the protein population should not change drastically. However, the overlap between the two experiments leads to 1343 proteins in common which represents a 60% overlap. That is a bit weak. Is this because of the stochastic nature of tandem mass spectra acquisition leading to different peptides being sequenced between each run or are we truly seeing two different protein population. A simple way to assess that would be to consider for the Venn diagram only the proteins identified with 2 peptides and more. Another option would be to cluster the identifications based on retention time charge and m/z and to match the identifications between runs (only if ProteinPilot can do it).

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Version 1

Reviewer Report 16 June 2015

<https://doi.org/10.5256/f1000research.7008.r9056>

© 2015 Fricker L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lloyd D. Fricker

Department of Molecular Pharmacology, Albert Einstein College of Medicine of Yeshiva University,

Bronx, NY, USA

In reading the manuscript, I initially thought that the authors had demonstrated that their modification of a technique for proteomics was a modest improvement over previously published methods. However, when I opened the supplemental datasets, I found some issues with the data. For example, in the FASP A10k data, there are 1199 proteins listed as “identified” but 401 of these were identified from a single peptide, and 5 identified from “0” peptides. How can a protein be identified without at least one confident peptide? Some laboratories that perform proteomic studies require 2 or more distinct peptides for a positive identification; some require 5 or more. If the latter, only 332 of the 1199 proteins would make the cut-off.

In addition, there are columns in the supplemental dataset with no data (log precursor, fraction of the total signal, Apex retention time). What’s the point of including a column with no data? The supplemental datasets are not helpful in understanding the results, and should be modified.

The main issue is whether the increased number of “identified” proteins reflects a real improvement in the technique, or just more noise in the system that gives tentative identifications that may not be real. In looking through the supplemental data, it was not clear to me that this is the case. At a minimum, the supplement needs to be edited and proteins based on “0” peptides should be removed. This will change many of the other figures/tables, which should only reflect proteins identified from 1 or more peptide! (2 would be preferable, but “0” is not acceptable)

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 14 Sep 2015

David Pellerin, Université de Sherbrooke, Sherbrooke, Canada

Reply from authors

Thank you very much for your valuable input on our paper and datasets. Below you will find detailed explanations to all of your comments. We would also like to apologize for the delay in making this response. Unfortunately, we were waiting for the comments of the second reviewer before making changes to our paper and we are still waiting for them.

First, the datasets have been restructured and modified according to the proteomic publication criterion of the *Molecular and Cellular Proteomics* journal (MCP publication standards http://www.mcponline.org/site/misc/Required_Manuscript_Guidelines_06-01-15.pdf). We are now providing new valuable information in the datasets, which may help to understand the purpose of our work. Precisely, protein identification reports, peptide summaries and peptide to spectra identification reports as well as the False Discovery Rate (FRD) tables (FDR analysis for protein, peptide and peptide to spectra match) are now presented in our supporting information. We also corrected the columns with no data, which have been inadvertently introduced in the dataset. In fact, there was probably a

problem with data transfer from our analysis software. Everything is now corrected.

Second, we apologize for the possible lack of clarity regarding the type of data analysis we performed in this study, which may have shed some confusion on the few proteins listed as identified with “0 peptide”. ProteinPilot unlike Mascot and Sequest reports 2 types of FDR score. The standard FDR analysis, termed global FDR, with a cut-off set at 99% (1% of the list would be accepted as a bad identification) is what we have used because we were not looking at particular and precise protein within the list. We were indeed performing shotgun and large-scale analysis for looking at the general tendency of protein identification between the different methods. However, when you look at a particular protein within the data set, it is recommended to use the local FDR (with a cut-off set at 95%) analysis that use a non-linear fitting model as published in [Tang *et al.* \(2008\)](#).

Also we would like to provide additional information regarding the different FDR analysis using Proteomics System Performance Evaluation Pipeline Software (PSPEP) from AB SCIEX: About Global and Local False Discovery Rates

One of the main benefits of the PSPEP analysis is that it provides error estimation using a variety of methods and for a range of thresholds. FDRs can be divided into two categories – local and global. The local FDR estimates the “local” error rate around a given identification, which indicates the likelihood that that the specific identification is incorrect. The global FDR estimates the error rate of the whole “global” set of answers defined by a threshold value. That is, the global FDR tells you the likely error rate of the entire set of identifications with scores as good as or better than the threshold.

PSPEP calculates the global FDR in two different ways:

- Using the conventional q-value approach
- Using a fitted curve as a smoothing tool, to provide a more reproducible estimate

For global FDRs, evaluate the quality of the global FDR from fit using the Nonlinear Fitting plot. If the fit is good, the global FDR from fit is probably more accurate than the conventional global FDR. But if the fit is poor (or fails), the global conventional FDR is clearly better. It is important to understand that in the set with 5% global FDR, there are proteins that may have local error rates much higher than 5%. The type of FDR you use and the threshold setting you select, depend on your purpose. In general, local error rates are more meaningful, but global error rates can generally be determined with higher precision, so global FDRs may be better for some comparison purposes.

Even though using the results obtained with the local FDR setting would change the total number of proteins in each group, this change should be coherent with the data quality of each group. Below is an example showing that changing the FDR cut-off does not really change the “tendency” and the “general meaning” of the data.

[Figure 1 \(link\)](#)

Concerning the “0 peptide-based” protein identification, these results are based on unique and over 95% confidence ProteinPilot peptide specific score (see below for explanation). One should also look at protein coverage based on confidence since ProteinPilot reports three score-based coverage (95%, 50% peptide confidence-based coverage, as well as

overall coverage).

Example of a "0 peptide" identified protein.

The 95% local peptide FDR for FASP A10k indicates that 5363 peptides are valuable (N=5363). This corresponds to a ProteinPilot score of 94%. In the other hand, the 95% global peptide FDR indicates that 5755 peptides are valuable (N=5755). This corresponds to a protein pilot score of 89%. If we look at protein # 1174, which falls within the global protein 99% FDR cut-off but presents a "0 peptide" feature, and look at its corresponding peptide we see that the ProteinPilot score of this corresponding peptide is 94.92. This score is below both local and global FDR cut-off, meaning that we have actually identified the protein with one good peptide. Also, if you look at the "% coverage (50%)" column you can see that we can identify 2,4% of the protein; it correspond to the peptide itself. However, doing this exercise is exactly what should not be done with protein that fall only within the 99% global protein FDR cut-off and it is advised to use the 95% local protein FDR cut-off if the intention is to look at particular/precise protein identification. Within the 95% local protein FDR cut-off you will find most of the proteins with one high quality peptide, redundant acquisition of that peptide or more than one peptide.

[Figure 2 \(Link\)](#)

[Figure 3 \(Link\)](#)

Competing Interests: No competing interests to disclose.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research