

Sample Preparation by Easy Extraction and Digestion (SPEED) - A Universal, Rapid, and Detergent-free Protocol for Proteomics Based on Acid Extraction

Authors

Joerg Doellinger, Andy Schneider, Marcell Hoeller, and Peter Lasch

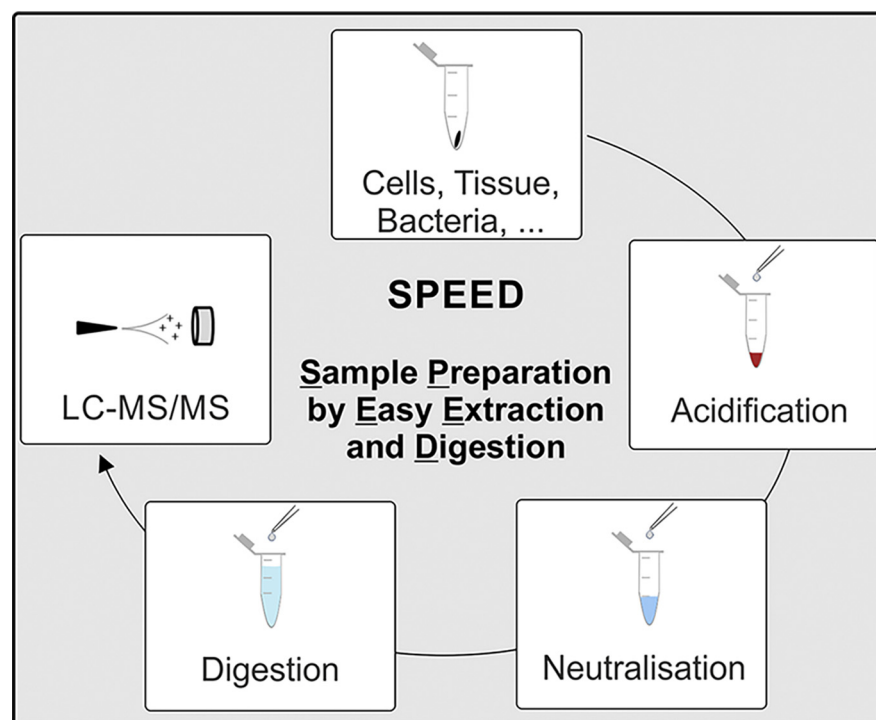
Correspondence

Doellingerj@rki.de

In Brief

Sample Preparation by Easy Extraction and Digestion (SPEED) is a universal method for peptide generation from various sources. Because of its detergent-free chemistry, SPEED offers some inherent benefits for bottom-up proteomics sample preparation, including improved quantitative reproducibility, simple applicability and enhanced proteome coverage for lysis-resistant sample types as well as very rapid sample processing. SPEED has high automatization potential and can contribute to simplify and standardize sample preparation and thus enhance reproducibility in proteomics.

Graphical Abstract



Highlights

- Universal and detergent-free proteomic sample preparation.
- Based on three simple mandatory steps (acidification, neutralization, digestion).
- Enhances proteome coverage especially for challenging samples.
- Improves quantitative reproducibility compared with ISD-Urea, FASP and SP3.

Sample Preparation by Easy Extraction and Digestion (SPEED) - A Universal, Rapid, and Detergent-free Protocol for Proteomics Based on Acid Extraction*[§]

Joerg Doellinger[‡], Andy Schneider, Marcell Hoeller, and Peter Lasch

The main challenge of bottom-up proteomic sample preparation is to extract proteomes in a manner that enables efficient protein digestion for subsequent mass spectrometric analysis. Today's sample preparation strategies are commonly conceptualized around the removal of detergents, which are essential for extraction but strongly interfere with digestion and LC-MS. These multi-step preparations contribute to a lack of reproducibility as they are prone to losses, biases and contaminations, while being time-consuming and labor-intensive. We report a detergent-free method, named Sample Preparation by Easy Extraction and Digestion (SPEED), which consists of three mandatory steps, acidification, neutralization and digestion. SPEED is a universal method for peptide generation from various sources and is easily applicable even for lysis-resistant sample types as pure trifluoroacetic acid (TFA) is used for highly efficient protein extraction by complete sample dissolution. The protocol is highly reproducible, virtually loss-less, enables very rapid sample processing and is superior to the detergent/chaotropic agent-based methods FASP, ISD-Urea and SP3 for quantitative proteomics. SPEED holds the potential to dramatically simplify and standardize sample preparation while improving the depth of proteome coverage especially for challenging samples. *Molecular & Cellular Proteomics* 19: 209–222, 2020. DOI: 10.1074/mcp.TIR119.001616.

The majority of mass spectrometry-based proteome studies are currently performed using a bottom-up approach, which relies on the digestion of proteins into smaller peptides (1). Different sample preparation methods have been developed aiming at enabling comprehensive and reproducible generation of peptides from proteomes extracted from a large variety of sample types. Current protocols employ detergents, e.g. sodium dodecyl sulfate (SDS)¹, or chaotropic agents, such as urea, for protein extraction and support sample lysis by physical disruption methods, such as heat, ultra-sonication

or grinding (2–4). As many extraction reagents inhibit enzymatic digestion of proteins and are incompatible with LC-MS/MS, the idea behind most sample preparation methods is to remove interfering substances before digestion either by filtration (FASP) (5), precipitation (on-pellet digestion, STrap) (6, 7) or bead-based purification (SP3) (8). However, enhancing lysis by use of physical disruption methods and subsequent protein purification requires additional sample handling steps, which are associated with attendant losses, biases and possible contaminations, while being time-consuming and labor-intensive.

The main exception is in-solution digestion (ISD) of proteins, which is based on the extraction of proteins using either digestion-compatible and often acid-labile surfactants or more frequently high urea concentrations and subsequent dilution into a concentration range which is not inhibiting tryptic digestion anymore (9–14). Urea-based ISD (ISD-Urea) benefits greatly from the reduced number of individual steps needed for sample preparation and is known to be robust, highly reproducible and easy to perform. However, ISD-Urea is not a universal method for bottom-up proteomics as urea contrasts with strong detergents, such as SDS, ineffective for extracting proteins from hard-to-lyse samples, such as tissues or Gram-positive bacteria. Further, urea is known to attach artificial modifications to proteins, known as carbamylation (15). This prevents samples from being lysed at elevated temperatures, which would be beneficial for proteome extraction especially from challenging samples. The limitations of in-solution digestion arise from the compromise between efficient lysis and low interference with enzymatic activity.

In this study we present a new sample preparation method, which overcomes the limitations of ISD but preserves its straight-forward approach. The new method is termed sample preparation by easy extraction and digestion (SPEED) and consists of three mandatory steps, namely acidification, neu-

From the Robert Koch-Institute, Centre for Biological Threats and Special Pathogens, Proteomics and Spectroscopy (ZBS6), Berlin, Germany

Received June 11, 2019, and in revised form, November 1, 2019

Published, MCP Papers in Press, November 21, 2019, DOI 10.1074/mcp.TIR119.001616

This is an open access article under the CC BY license.

tralization and digestion. SPEED uses neither detergents nor chaotropic agents for protein extraction but pure trifluoroacetic acid. The method is robust, highly-reproducible, well-suited even for lysis-resistant sample types, inexpensive, requires low hands-on time and can easily be performed by non-experts without the need for special equipment.

EXPERIMENTAL PROCEDURES

Sample Materials—Tryptic Soy Agar (TSA) ReadyPlates™ (Merck, Darmstadt, Germany) were inoculated with *E. coli* K-12 (DSM 3871), *S. aureus* (DSM 4910) or *B. cereus* (ATCC® 10987) and incubated at 37 °C overnight. Cells were harvested using an inoculating loop and washed in 2 × 1 ml phosphate-buffered saline (PBS) for 5 min at 4000 × *g* and 4 °C. Cells were aliquoted, again pelleted and kept at –80 °C until lysis.

HeLa cells (ATCC® CCL-2™) were cultivated in DMEM supplemented with 10% FCS and 2 mM L-Glutamine at 37 °C and harvested at 90% confluency by scraping. Cells were washed in 2 × 2 ml phosphate-buffered saline (PBS) for 8 min at 400 × *g* and 4 °C, aliquoted, again pelleted and kept at –80 °C until lysis.

C57BL/6 mouse lung and liver were obtained from an animal of the central experimental animal facility (MF3, Robert Koch Institute, Berlin, Germany) which was killed in experiments approved by the local authority. Tissue samples were washed 3 x by dipping into 10 ml ice-cold PBS, cut into equal slices, aliquoted, flash frozen with liquid nitrogen and kept at –80 °C until lysis. After addition of lysis buffer (FASP, SP3, Urea-ISD) mouse lung samples were transferred to Lysing Matrix D Tubes (MP Biomedicals, Santa Ana, CA) and subjected to grinding.

Lyophilized microbiome standard (20 Strain Even Mix Whole Cell Material, ATCC® MSA2002™) was resuspended in 1 ml PBS, divided into 15 aliquots and cells were pelleted for 15 min at 5000 × *g* and 4 °C.

Sample Preparation by Easy Extraction and Digestion (SPEED)—Samples were resuspended in trifluoroacetic acid (TFA) (Uvasol® for spectroscopy, Merck) (sample/TFA 1:4 (v/v) or 10 µl for the *E. coli* 1 µg experiment) and incubated at room temperature for 2 (*E. coli*, *B. cereus*, *S. aureus*, HeLa, Microbiome Standard) or 10 (mouse lung) min. *B. cereus*, *S. aureus* and microbiome samples were further irradiated for 10 s at 800 W using a microwave oven. Samples were neutralized with 2 M TrisBase using 10 x volume of TFA and further incubated at 95 °C for 5 min after adding Tris(2-carboxyethyl)phosphine (TCEP) to a final concentration of 10 mM and 2-Chloroacetamide (CAA) to a final concentration of 40 mM. Protein concentrations were determined by turbidity measurements at 360 nm (1 AU = 0.79 µg/µl) using GENESYS™ 10S UV-Vis Spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and adjusted to 0.25 µg/µl using a 10:1 (v/v) mixture of 2 M TrisBase and TFA and then diluted 1:5 with water. Digestion was carried out for 20 h at 37 °C using Trypsin Gold, Mass Spectrometry Grade (Promega, Fitchburg, WI) at a protein/enzyme ratio of 50:1. A more detailed description of the usage of SPEED is provided as a lab protocol in the supplementary materials.

CAUTION—Trifluoroacetic acid is highly corrosive and must be handled using appropriate Personal Protective Equipment!!!

Preparation of SPEED-derived Samples for LC-MS/MS Measurements Using Online Desalting on Trap Columns—HeLa and *E. coli* cells were prepared using SPEED as previously described except that LC-MS grade TFA and water was used (Thermo Fisher Scientific). Digested samples were acidified using TFA (final concentration 2%) and further filtrated on 200 µl StageTips packed with two MK 360 Micro-Quartz Fiber Filters (MUNKTELL & FILTRAK, Baerenstein, Germany) at 2000 × *g* for 2 min.

Filter-aided Sample Preparation by Easy Extraction and Digestion (fa-SPEED)—HeLa cells were resuspended in trifluoroacetic acid (TFA) (Uvasol® for spectroscopy, Merck) (sample/TFA 1:4 (v/v)) and incubated at room temperature for 2 min. Samples were neutralized with 2 M TrisBase using 10 x volume of TFA and further incubated at 95 °C for 5 min after adding Tris(2-carboxyethyl)phosphine (TCEP) to a final concentration of 10 mM and 2-Chloroacetamide (CAA) to a final concentration of 40 mM. Protein concentrations were determined by turbidity measurements at 360 nm (1 AU = 0.79 µg/µl) using GENESYS™ 10S UV-Vis Spectrophotometer (Thermo Fisher Scientific). 50 µg proteins were diluted to 40 µl using a 10:1 (v/v) mixture of 2 M TrisBase and TFA, mixed with 160 µl acetone and incubated for 2 min at RT. Proteins were captured on Ultrafree®-MC (0.5 ml) centrifugal devices, 0.2 µm, PTFE (Merck) at 5000 × *g* for 2 min. The samples were washed successively with 200 µl 80% acetone, acetone and n-pentane at 5000 × *g* for 2 min each. Afterward 40 µl digestion buffer (50 mM ammonium bicarbonate or 1:10 Rapid Digest buffer (Promega)) containing trypsin (1:100 Trypsin Gold, Mass Spectrometry Grade (Promega) or 1:10 Rapid Digestion Trypsin (Promega)) was added and samples were incubated at 37 °C (20h) or 70 °C (15 and 60 min). Peptides were eluted at 5000 × *g* for 2 min and acidified using formic acid before LC-MS injection. A more detailed description of the usage of fa-SPEED is provided as a lab protocol in the supplementary materials.

Filter-aided Sample Preparation (FASP)—Samples were suspended in 4% SDS, 100 mM Tris/HCl, 100 mM DTT, pH 7.6 (sample/buffer 1:10 (v/v) or 10 µl for the *E. coli* 1 µg experiment), incubated at 95 °C for 5 min and further sonicated for 10 (*E. coli*, HeLa) or 15 (*B. cereus*, mouse lung tissue) cycles a 30 s at high intensity level and 4 °C using Bioruptor®Plus (Diagenode, Liege, Belgium). Samples were clarified by centrifugation at 16,000 × *g* for 5 min und processed using Microcon-30kDa Centrifugal Filter Units (Merck) according to the Filter-aided Sample Preparation (FASP) protocol of Wisniewski *et al.* (5) with proteins being digested for 20 h at 37 °C using Trypsin at a protein/enzyme ratio of 50:1.

Single-Pot Solid-Phase-enhanced Sample Preparation (SP3)—Samples were suspended in 1% SDS, 1x cOmplete Protease Inhibitor Mixture (Roche, Basel, Switzerland), 50 mM HEPES buffer, pH 8.5 (sample/buffer 1:10 (v/v) or 10 µl for the *E. coli* 1 µg experiment), incubated at 95 °C for 5 min and further sonicated for 10 (*E. coli*, HeLa) or 15 (*B. cereus*, mouse lung tissue) cycles a 30 s at high intensity level and 4 °C using Bioruptor®Plus. Samples were further processed according to the Single-Pot Solid-Phase-enhanced Sample Preparation (SP3) method of Hughes *et al.* (8), except that proteins were bound to the paramagnetic beads at 70% ACN without acidification as prescribed by Sielaff *et al.* (3). Digestion was carried out for 20 h at 37 °C using Trypsin at a protein/enzyme ratio of 50:1.

Urea-based In-solution Digestion (Urea-ISD)—Samples were suspended in 8 M urea, 50 mM Tris-HCl, 5 mM DTT, pH 8 (sample/buffer 1:10 (v/v) or 10 µl for the *E. coli* 1 µg experiment) and sonicated for 10 (*E. coli*, HeLa) or 15 (*B. cereus*, mouse lung tissue) cycles a 30 s at high intensity level and 4 °C using Bioruptor®Plus. Samples were further incubated for 1 h at 37 °C and clarified by centrifugation at 16,000 × *g* for 5 min. IAA was added to a final concentration of 15 mM

¹ The abbreviations used are: SDC, sodium deoxycholate; CAA, 2-Chloroacetamide; DTT, Dithiothreitol; FASP, filter-aided sample preparation; FDR, false discovery rate; HCD, higher-energy c-trap dissociation; IAA, iodoacetamide; ISD, in-solution digestion; iST, instagetip sample preparation; NCE, normalized collision energy; PSM, peptide spectrum match; SP3, single-pot solid-phase-enhanced sample preparation; SPEED, sample preparation by easy extraction and digestion; Strap, suspension trapping; TCEP, tris(2-carboxyethyl)phosphine; THSD, Tukey's honestly significant difference.

and samples were alkylated for 30 min at room temperature in the dark. Urea was diluted with 50 mM Tris-HCl (pH 8) to 1 M, Trypsin was added at a protein/enzyme ratio of 50:1 and proteins were digested for 20 h at 37 °C.

In-StageTip Sample Preparation (iST)—Microbiome standard was prepared using the iST Kit according to manufacturer's instructions (Preomics, Munich, Germany) in triplicates. Bacterial cell lysis was supported by sonication for 15 cycles a 30 s at high intensity level and 4 °C using Bioruptor®Plus.

Suspension Trapping (STrap)—Microbiome standard was suspended in 5% SDS, 20 mM DTT, 50 mM Tris/HCl buffer, pH 7.6 (sample/buffer 1:10 (v/v)), incubated at 95 °C for 10 min and further sonicated for 15 cycles a 30 s at high intensity level and 4 °C using Bioruptor®Plus. Samples were further processed according to the Suspension trapping (STrap) protocol (6).

Measurements of Tryptophan Fluorescence—Protein concentrations were determined by measuring the tryptophan fluorescence at an emission wavelength of 350 nm using 295 nm for excitation with an Infinite® M1000 PRO microplate reader (Tecan, Maennedorf, Switzerland) (16). The tryptophan content of each sample was determined using a standard curve ranging from 0.1–0.9 µg tryptophan and assuming a tryptophan weight content of 1.3% in the samples. Twenty micrograms protein of each sample type was further digested except for the *E. coli* experiment with low starting amount (*E. coli* 1 µg). Therefore, protein content of a cell suspension was determined after TFA-based lysis and cells from the suspension volume equivalent to 1 µg protein were aliquoted and pelleted.

The calibration curves for the turbidity measurements (Fig. 5) were determined by dilution of SPEED-derived samples (*E. coli*, *B. cereus*, HeLa, Mouse Liver) in 1.7 M TrisBase, 8% TFA, 1.7% SDS. Samples were incubated for 5 min at 95 °C to solubilize proteins. Afterward, fluorescence was measured at an emission wavelength of 350 nm using 295 nm for excitation. A tryptophan standard curve in the range of 0.1–0.9 µg solubilized in 1.7 M TrisBase, 8% TFA, 1.7% SDS was measured in triplicates and was used for protein content determination assuming a tryptophan weight content of 1.3% in the samples.

Turbidity Measurements—Turbidity of SPEED-lysates or MacFarland standard series (bioMérieux, Marcy-l'Étoile, France) was measured at 360 nm using either a GENESYS™ 10S or an Implen NP80 (Implen, Munich, Germany) UV-Vis Spectrophotometer. If the initial protein concentration was above 1 µg/µl, samples were diluted with a 10:1 (v/v) mixture of 2 M TrisBase and TFA. For experiments analyzing the digestion progress by real-time monitoring, an Infinite® M1000 PRO microplate reader was used. The microplate reader was tempered to 37 °C and turbidity was measured at 360 nm every 5 min.

Peptide Desalting—Peptides generated using FASP, SPEED, and Urea-ISD were desalted using 200 µl StageTips packed with three Empore™ SPE Disks C18 (3 M Purification, Inc., Lexington, KY) according to Rappsilber *et al.* (17) and concentrated using a vacuum concentrator. Samples were resuspended in 20 µl 0.1% formic acid (FA) and peptides were quantified by measuring the absorbance at 280 nm using a Nanodrop 1000 (Thermo Fisher Scientific).

High-pH C18-StageTip-based Peptide Fractionation—Peptide solutions were acidified after SPEED-based processing using TFA and loaded onto six Empore™ SPE Disks C18 packed in 200 µl pipette tips, which were activated using methanol. Peptides were washed with 0.1% TFA, 100 mM ammonium formate (pH = 10) and 20 mM ammonium formate (pH = 10) consecutively. Afterward the peptides were eluted sequentially in 5/7.5/10/12.5/15/17.5/20/50 (v/v) ACN in 20 mM ammonium formate (pH = 10) and dried down using a vacuum concentrator. Samples were resuspended in 20 µl 0.1% FA and peptides were quantified by measuring the absorbance at 280 nm using a Nanodrop 1000.

Phosphopeptide Enrichment—SPEED was used to digest HeLa cells (100 µg) before phosphopeptide enrichment using the high-sensitivity EasyPhos workflow (12). Briefly, 300 µl of the digested samples were mixed with 400 µl isopropanol, 100 µl 48% (v/v) TFA and 8 mM KH₂PO₄ and 5 mg Titanium dioxide (TiO₂) beads (GL Sciences, Tokyo, Japan). Samples were incubated at 40 °C for 5 min and subsequently washed five times using 5% (v/v) TFA/60% (v/v) isopropanol. Beads were removed in a transfer buffer of 0.1% (v/v) TFA/60% (v/v) isopropanol using 200 µl StageTips packed with two MK 360 discs. Phosphopeptides were eluted twice in 30 µl 40% ACN and 4% NH₄OH and purified using SDB-RPS StageTips (3 M Purification, Inc.) as described elsewhere (12).

Liquid Chromatography and Mass Spectrometry—Peptides were analyzed on an EASY-nanoLC 1200 (Thermo Fisher Scientific) coupled online to a Q Exactive™ Plus mass spectrometer (Thermo Fisher Scientific). 1 µg peptides were separated on a 50 cm Acclaim™ PepMap™ column (75 µm i.d., 100 Å C18, 2 µm; Thermo Fisher Scientific) using a linear 120 (*E. coli*, *B. cereus*, high-pH fractions), 180 (HeLa, mouse lung tissue) or 240 (microbiome standard) min gradient of 3 to 28% acetonitrile in 0.1% formic acid at 200 nL/min flow rate. Column temperature was kept at 40 °C using a butterfly heater (Phoenix S&T, Chester, PA).

Samples for evaluating online desalting, filter-aided SPEED (fa-SPEED) and phosphopeptide analysis were loaded on a Acclaim™ PepMap™ trap column (20 mm × 75 µm i.d., 100 Å, C18, 3 µm; Thermo Fisher Scientific) at a flow rate of 3 µl/min and were subsequently separated on 200 cm µPAC™ column (PharmaFluidics, Ghent, Belgium) using a 160 min (online desalting) or 210 min gradient (filter-aided SPEED, phosphoproteomics) of 3 to 28% acetonitrile in 0.1% formic acid at 300 nL/min flow rate. Column temperature was kept at 50 °C using a butterfly heater (Phoenix S&T, Chester, PA).

The Q Exactive™ Plus was operated in a data-dependent manner in the *m/z* range of 300–1650. Full scan spectra were recorded with a resolution of 70,000 using an automatic gain control (AGC) target value of 3×10^6 with a maximum injection time of 20 ms. Up to the 10 most intense 2^+ – 5^+ charged ions were selected for higher-energy c-trap dissociation (HCD) with a normalized collision energy (NCE) of 25%. Fragment spectra were recorded at an isolation width of 2 Th and a resolution of 17,500 at 200 *m/z* using an AGC target value of 1×10^5 with a maximum injection time of 50 ms. The minimum MS² target value was set to 1×10^4 . Once fragmented, peaks were dynamically excluded from precursor selection for 30 s within a 10 ppm window. Peptides were ionized using electrospray with a stainless-steel emitter, I.D. 30 µm, (Proxeon, Odense, Denmark) at a spray voltage of 2.0 kV and a heated capillary temperature of 275 °C.

For the analysis of phosphorylated peptides the maximum injection times were increased to 200 ms (full scan) and 100 ms (MS²) as well as the minimum MS² target value to 2×10^4 . Further, the isolation width was decreased to 1.6 Th and precursor ion charges were restricted to 2–4. All other MS parameters were unaltered.

Data Analysis—Mass spectra were analyzed using MaxQuant (Version 1.5.1.8 and 1.6.1.0) (18). At first, parent ion masses were recalibrated using the “software lock mass” option before the MS² spectra were searched using the Andromeda algorithm against sequences from the complete proteomes of either *homo sapiens* (UP000005640, 92578 sequences, 15/7/16), *E. coli* K-12 (UP000000625, 4315 sequences, downloaded 26/10/16), *B. cereus* strain ATCC 10987 (UP000002527, 5821 sequences, downloaded 8/3/17), *S. aureus* strain NCTC 8325 (UP000008816, 2889 sequences, downloaded 4/10/18) or *Mus musculus* (UP000000589, 58430 sequences, downloaded 10/1/17), which were downloaded from UniProt. Proteome entries of bacterial species from the microbiome standard were downloaded from UniProt according to their corresponding ATCC® numbers (62989 sequences, downloaded 31/7/18). If no entry was

available, the reference proteome of the respective species was used. Spectra were searched with a tolerance of 4.5 ppm in MS¹ and 20 ppm in HCD MS² mode, strict trypsin specificity (KR not P) and allowing up to two missed cleavage sites. Cysteine carbamidomethylation was set as a fixed modification and methionine oxidation as well as N-terminal acetylation of proteins as variable modifications. For phosphopeptide analysis phosphorylation (S,T,Y) was added as an additional variable modifications. The false discovery rate (FDR) was set to 1% for peptide and protein identifications. Dependent peptides with previously unconsidered modifications were identified using 1% FDR as well. Identifications were transferred between samples using the 'match between run' option within a match window of 0.7 min and an alignment window of 20 min. Protein intensities for label-free quantification (LFQ) were calculated separately for each sample preparation method. The analysis was once repeated for the method comparison experiments allowing semi-specific tryptic digestion at the peptide N termini.

Analysis of the MaxQuant results was done in Perseus (Version 1.5.0.31 and 1.6.1.1) (19). At first, reverse hits, contaminants and proteins only identified by a modification site were removed. Afterward protein related parameters (identified proteins, quantified proteins, Pearson correlation coefficients, coefficients of variation), peptide related parameters (identified peptides, quantified peptides, missed cleavage sites, ragged-N peptides, modified peptides (C, oxM)) and dependent peptide related parameters (distribution of sample preparation-related peptide modifications) were extracted from the respective .txt files. For comparison of the different sample preparation protocols, the number of ragged-N peptides, modified peptides (C, oxM) and dependent peptides were normalized to the number of unmodified tryptic peptide identifications of the same run. Only peptides and proteins without missing values were considered for quantification. Significant protein expression differences between mouse lung samples were identified using FDR-adjusted *p* values from an ANOVA test with a permutation-based FDR of 0.01 and 250 randomizations. Significant differences between sample preparation methods were further analyzed by calculating the Tukey's honestly significant difference (THSD) using an FDR of 0.01. THSD were further used for gene ontology (GO) enrichment analysis using a Benjamini-Hochberg FDR threshold of 0.01. Quantification of bacterial species was based on summed iBAQ-intensities of proteins with at least one species-unique peptide. Species abundances were further log₂-transformed, normalized to the mean and differentially extracted bacteria were identified by a *t* test with a permutation-based FDR of 0.01. Protein copy numbers in the fractionated *E. coli* sample were calculated based on the total protein approach using the protein intensity values (20).

Open search analysis of the mouse lung data using wide precursor mass windows (500 Da) was done using MSFragger GUI (v6.0) with a precursor mass tolerance of 500 Da and fragment mass tolerance of 20 ppm, strict trypsin specificity (KR not P) and allowing up to one missed cleavage site. Peptide length was set to 7–50 amino acids within the mass range of 500–5000 Da.(21). Peptide and protein identifications were filtered using a false discovery rate (FDR) of 1%.

Experimental Design and Statistical Rationale—All combinations of sample types (HeLa, *E. coli*, *E. coli*-1 µg, *B. cereus*, mouse lung tissue, ATCC® Microbiome Standard) and preparation methods (FASP, ISD-Urea, SP3, SPEED, STrap, iST) for the sample preparation benchmarking experiment (total number of samples = 60) and the microbiome preparation comparison (total number of samples = 9) were prepared and measured in triplicates. LC-MS measurements were shuffled in such a way, that replicates of the same sample and preparation type were not measured consecutively. Sample preparation methods were compared using parameters calculated from these

triplicate analyses (Figs. 3 and 5). Further details are described in the section Data Analysis.

RESULTS

Development of SPEED—The main challenge during sample preparation is to extract all proteins in a manner that enables efficient digestion into peptides and is compatible with subsequent mass spectrometric analysis. A universal bottom-up proteomics sample preparation method would ideally consist of a single extraction step, which is independent of the sample type, and allows subsequent proteolytic digestion without the need for protein purification. To develop such a one-step universal protein extraction method, we decided to leave the track of common lysis buffer chemistry and ended up using pure trifluoroacetic acid (TFA).

TFA is a strong acid ($pK_a = 0.2$) and an excellent solvent for proteins (22). Although high concentrations of formic acid and triflic acid have been used before for the solubilization of precipitated protein pellets to facilitate the analysis of subproteomes, strong acids have not been used as a substitute for detergents during the lysis of biological material yet (14, 23). We found that pure TFA can dissolve cells and tissues within few minutes at room temperature forming clear lysates. Viscosity of the lysates is low as DNA is degraded rapidly after acidification. Analysis of *E. coli* cells incubated with TFA for different times between 1 and 60 min revealed, that TFA does neither hydrolyze peptide bonds nor modifies amino acid residues (supplemental Fig. S1). Protein intensities of samples with different TFA incubation times had excellent Pearson coefficients of at least 0.994 and numbers of protein and peptide identifications were unaffected (supplemental Fig. S1).

TFA lysates are prepared for proteolytic digestion by neutralization with a weak base, whose pK_a is equal to the optimal pH range of the protease of choice. In such a high capacity buffer system, the desired pH value is stable over a wide range of volume ratios between TFA and base (supplemental Fig. S2). Tris(hydroxymethyl)-aminomethan (Tris) ($pK_a = 8.1$) was chosen to neutralize TFA lysates for tryptic digestion. After neutralization, lysates become slightly turbid as proteins precipitate and form fine homogeneous particles, which don't sediment because of the high salt content of the sample. As neutralization is an exothermic reaction, samples instantly heat up to 70–80 °C. Heating can be exploited to shorten the incubation time needed for reduction and alkylation of disulfide bonds after instant addition of Tris(2-carboxyethyl)phosphine (TCEP) and 2-Chloroacetamide (CAA) down to 5 min as previously described (10).

The digestion progress of SPEED-derived lysates is associated with the complete disappearance of the protein particles and thus also the turbidity of the sample, as the resulting peptides are water soluble. Therefore, proteins in dispersions of fine homogeneous particles must be highly accessible for trypsin presumably resulting from the large surface area and

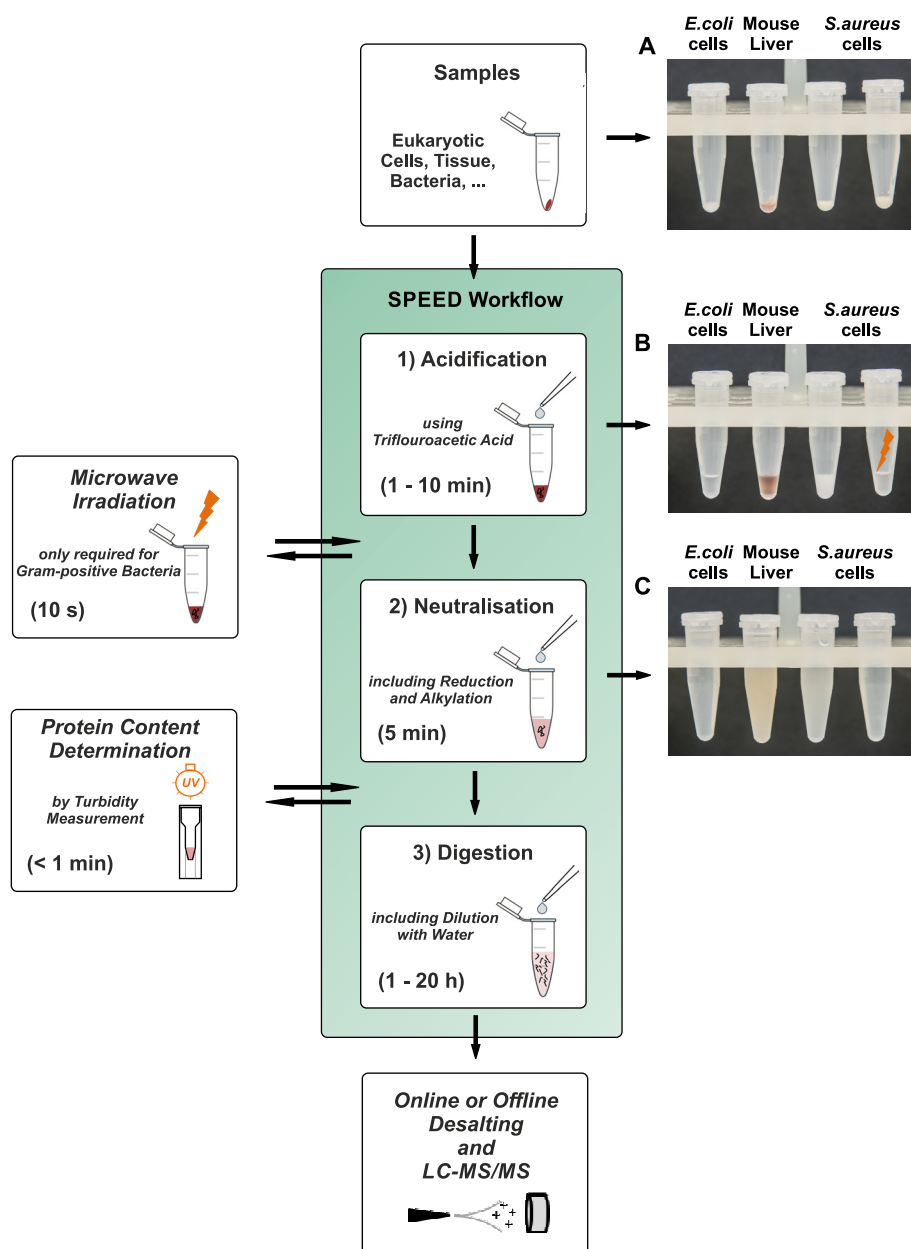


FIG. 1. **Workflow of Sample Preparation by Easy Extraction and Digestion (SPEED).** Sample processing using SPEED consists of three mandatory steps, namely acidification with pure trifluoroacetic acid (TFA) (1), neutralization with 2 M TrisBase (2) and digestion (3). Although TFA dissolves *E. coli* cells and mouse liver tissue at room temperature completely in a few minutes, lysis of Gram-positive bacteria *S. aureus* requires an additional step of microwave irradiation for 10 s. After neutralization the samples become slightly turbid, as proteins precipitate. Proteins are quantified in the resulting dispersions by turbidity measurement at 360 nm and subsequently digested by addition of trypsin. Afterward peptides are desalted either online or offline before LC-MS/MS measurements.

can be digested efficiently without the need for any purification steps. Optimal results were obtained after diluting the dispersion 3–10-fold with water to reduce the molarity of salts (supplemental Fig. S3). Trypsin digestion in dispersion was found to proceed like in-solution digestion. At 37 °C and overnight incubation (20 h), the optimal enzyme-to-protein ratio was found to vary between 1:20 and 1:100, depending on the protein concentration (supplemental Fig. S4).

The resulting protocol for *Sample Preparation by Easy Extraction and Digestion (SPEED)* consists of three mandatory steps, namely acidification, neutralization and digestion. The execution of SPEED is exemplified for *E. coli* and *S. aureus* cells as well as for mouse liver tissue in Fig. 1. TFA dissolves *E. coli* and mouse liver tissue completely within 1 and 10 min respectively, whereas acidification of Gram-positive bacteria *S. aureus* forms a cloudy lysate because of insoluble cell wall

components (Fig. 1B). Therefore, preparation of Gram-positive bacteria requires microwave irradiation of 10 s or incubation at 70 °C for 3 min before instant neutralization. Further, like detergent-based lysis buffers, TFA does not completely dissolve large collagenous fibers, e.g. from skin preparations. In this case, however, this behavior is not undesirable, as the high abundance of fiber proteins exceeds the abundance of all other proteins by far which would impede comprehensive analysis of tissue proteomes.

Complete sample dissolution in TFA with consequent neutralization is a purely chemical approach for protein extraction in contrast to the well-characterized detergent-based methods, which are usually assisted by physical disruption techniques. Therefore, the influence of acidification on the analysis of endogenous post-translational modifications (PTM) must be determined independently for each distinct modification. It is likely that some modifications, e.g. glycosylations, are destroyed during acidification whereas many other PTMs (phosphorylation (S,T,Y), methylation, acetylation, ubiquitination) are presumably stable at low pH. To obtain an initial assessment if SPEED could be suitable for PTM analyses, we choose to analyze phosphorylations, as they are currently the most-studied PTM in proteomics. Therefore, HeLa cells (100 µg) were prepared in triplicates without any optimization using a combination of SPEED and EasyPhos-based phosphopeptide enrichment and compared with the original high-sensitivity EasyPhos protocol using sodium deoxycholate (SDC) for lysis (12). SPEED enabled to identify more phosphorylation sites (S,T,Y) concerning either all sites (6412 *versus* 5923) or after filtering with a localization probability cutoff of 0.75 (4228 *versus* 3525). The ratios between modified amino acids were highly similar (S:T:Y = 0.847:0.148:0.047 (SPEED) *versus* 0.845:0.151:0.039 (SDC)). This initial data shows, that phosphoproteomics using SPEED is possible.

Protein content of SPEED-derived samples can be determined after adding SDS to a sample aliquot to solubilize the proteins before performing measurements of tryptophan fluorescence (16). However, the most straightforward approach for quantifying proteins in dispersions is by measuring the turbidity, which is an absolute measure for the protein content without the need for any sample manipulation. We found that turbidity measurements at 360 nm, which were calibrated using tryptophan fluorescence, provided consistent results for different sample types with a linear response for protein concentrations between 0.05 - 1 µg/µl (supplemental Fig. S5A). However, the turbidity values are instrument dependent (supplemental Fig. S5B). Calibration between different UV-spectrophotometers can be done using commercial MacFarland standards analogous to optical density (OD600) measurements in microbiology (supplemental Fig. S5C) according to the formula provided in supplemental Fig. S5D. Further digestion progress can be monitored in real-time using continuous turbidity measurements in a tempered microplate reader as peptides become soluble (supplemental Fig. S6).

During SPEED, proteins are digested in ~ 400 mM tris-trifluoroacetate, which is highly water-soluble and can easily be removed using reversed-phase C18-resins along with other salts as well as with monomeric nucleic acids and sugars, which are hydrolyzed during acidification. Therefore, crude samples can directly be subjected to nanoLC-MS using a trap column for online desalting. This procedure minimizes sample handling steps and enables very fast sample preparation from native samples to nanoLC injection.

The comparison of online desalting and Stage-tip-based purification of HeLa and *E. coli* cells using triplicate measurements revealed that the intensity of the total ion chromatograms and the peptide retention times of crude digests are indeed highly similar to offline desalted samples (Fig. 2A). The absence of ion suppression and retention time shifts proves that TFA in crude lysates was efficiently removed during sample trapping. Quantitative reproducibility was high and independent of the desalting procedure. The Pearson correlation coefficients between crude and StageTip-processed samples were > 0.98 for both samples and mean coefficients of variation were 10–11% (Fig. 2C–2E).

However, the numbers of protein and peptide identifications were reduced by 1–3% (protein) and 7–11% (peptides) in the crude lysates (Fig. 2B). The reason for this remains elusive. Probably, peptide trapping from high salt concentrations in the crude samples is slightly less efficient and needs optimization. Further, the exact peptide load is difficult to adjust between crude and offline desalted samples. The peptide amount loaded on the column was 2 µg for the StageTip-processed samples determined by UV-spectroscopy. For crude samples the volume corresponding to 2 µg protein determined by the protein turbidity measurements was injected as the actual peptide content cannot be determined by UV-spectroscopy. The injection volume of 10 µl was kept constant for all samples. Therefore, the resulting MS signal depended on the comparability of the protein turbidity and the peptide UV-absorption measurements as well as on the validity of the assumption that the protein to peptide conversion yield in the crude samples is 100%. Presumably these two assumptions are prone to errors, which led to slightly different loading amounts and in turn might have translated into different identification rates. The mean total ion current of the crude lysates was in any case slightly decreased by 6% (*E. coli*) and 9% (HeLa) compared with the StageTip-processed samples. Nevertheless, the data show that online desalting of crude samples delivers comparable results to offline desalting and represents a truly minimalist protocol, which enables sample preparation for bottom-up proteomics in just < 15 min hands-on time. A video presentation of using SPEED for tissue preparation is deposited at <https://zenodo.org/record/3552550>.

Minimization of the Total Processing Time—The straightforward nature of SPEED minimizes sampling handling and reduces hands-on time needed for sample preparation (Fig. 3A).

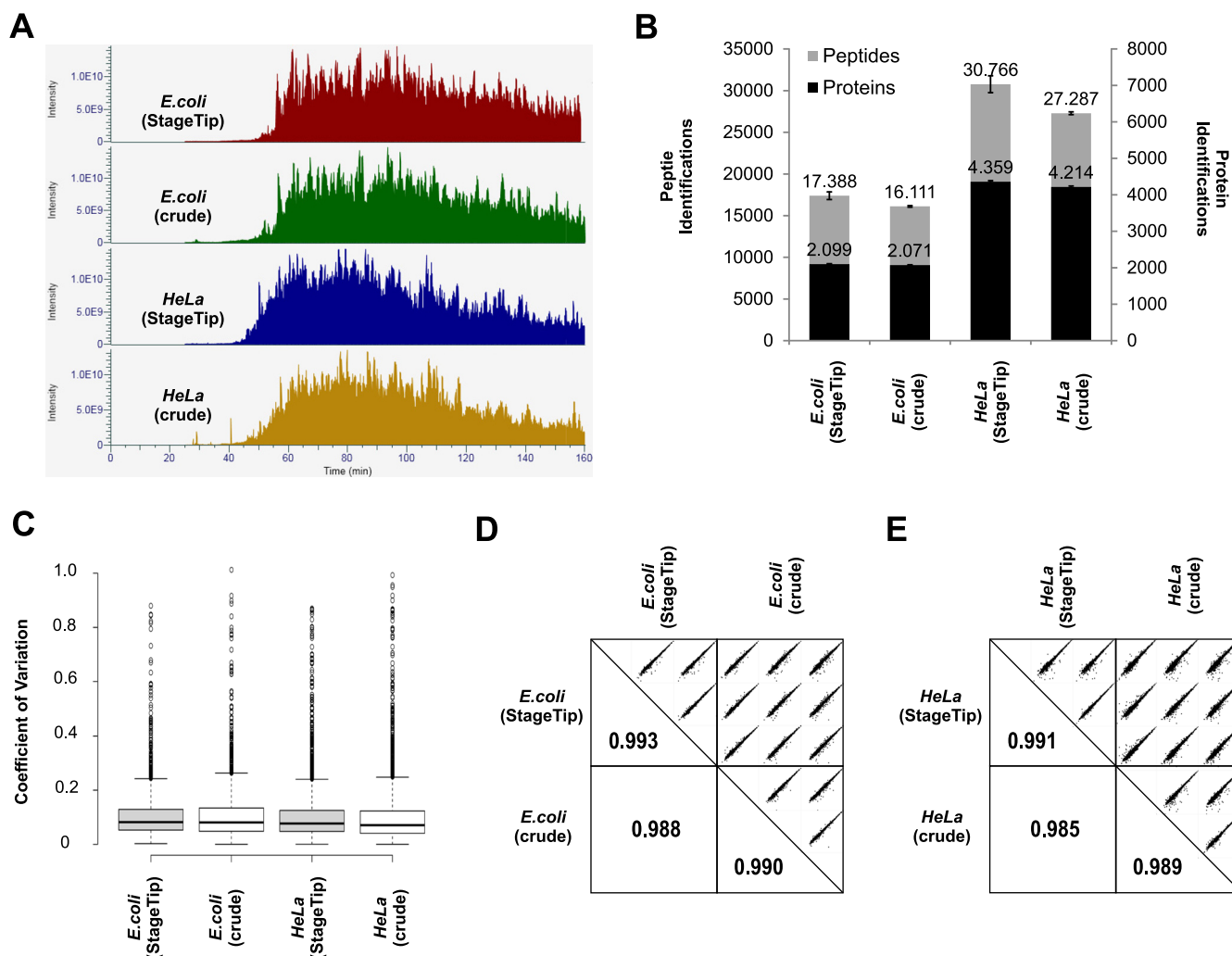


FIG. 2. Comparison of crude and StageTip-processed samples. The detergent-free approach of SPEED enables injection of crude samples without the need for any offline purification step into an LC-MS system equipped with a trap column for online desalting. *E. coli* and HeLa cells were used to compare the injection of crude samples and offline desalted samples using StageTips. Total ion chromatograms for the first replicate of each sample is shown in (A). Number of protein and peptide identifications are displayed in (B). The coefficients of variation and Pearson correlation coefficients of label-free protein quantification calculated from triplicate measurements are compared in (C), (D) and (E).

The potential for minimization of total processing time was evaluated using alternative approaches for digestion and peptide desalting, which mainly determine the total time required for SPEED. As offline desalting is slightly preferable considering identification rates and the completeness of protein digestion after strong shortening of the digestion process would most presumably benefit from increasing the protein concentration, a filter-aided SPEED (fa-SPEED) protocol was established. The principle is based on the capturing of aggregated proteins in a 0.2 μm spin filter. Aggregation of proteins is thereby enhanced by addition of acetone. The capturing of aggregated proteins on filters or micro-particles has already been demonstrated for detergent-based methods (6, 8, 24). It enables concentrating samples up to the desired concentration and offers the possibility to wash the proteins and thus to elute clean peptides.

The performance of filter-aided protocols with different digestion times was compared with the basic SPEED protocol for the triplicate analysis of HeLa cells. The results are summarized in Fig. 3. The original protocol ("SPEED 20 h, StageTip") provided most peptide and protein identifications and slightly outperformed the corresponding filter-aided protocol ("filter-aided SPEED 20 h") by 3% concerning proteins and 7% concerning peptides (Fig. 3B). The use of heat-stable trypsin ("filter-aided SPEED 15 min/1 h") reduced the numbers of protein and peptide identifications compared with the original protocol by 4% and 9–13% respectively. Interestingly, identification rates were quite similar after reduction of the digestion duration from 1 h to 15 min. A total processing time of only 37 min from cell pellet to LC-MS injection was enough to achieve 96% of the protein identifications of the original protocol. Quantitative reproducibility was excellent for

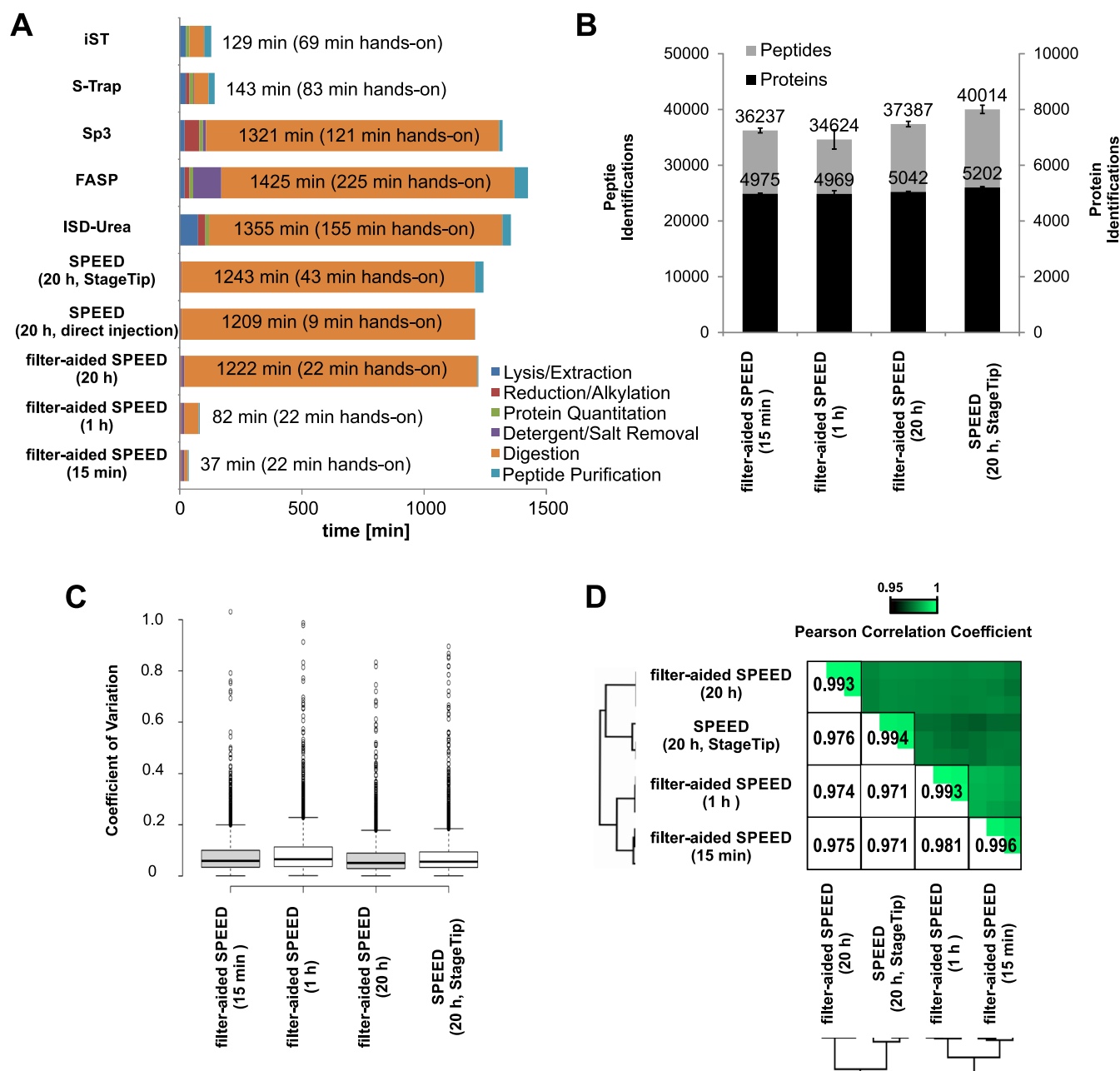


FIG. 3. Minimization of the total processing time. Filter-aided SPEED (fa-SPEED) protocols were established to reduce the total processing time. These protocols were compared with the basic SPEED protocol (“SPEED 20 h, StageTip”) for the triplicate analysis of HeLa cells by single-shot LC-MS. Hands-on and total processing times for different sample preparation protocols are compared in (A). The digestion duration of the SPEED protocols is indicated in parenthesis. Number of protein and peptide identifications are displayed in (B). The coefficients of variation and Pearson correlation coefficients of label-free protein quantification are compared in (C) and (D).

all samples and independent from the total processing time. Pearson correlation coefficients for the replicate preparations were > 0.99 within each protocol and the mean coefficients of variation were 6–7% (Fig. 3C–3D). Therefore, we conclude that the use of heat-stable trypsin in combination with filter-aided SPEED could be an attractive approach when time to results matter.

Comparison of SPEED with FASP, SP3 and Urea-ISD for the Analysis of Various Sample Types—The performance of SPEED was evaluated using triplicate preparations of different sample types compared with detergent-based methods (FASP, SP3) (5, 8) as well as to urea-based in solution digestion (ISD-Urea) (9). Sample selection consisted of an easy-to-lyse sample with medium complexity (*E. coli*) at different start-

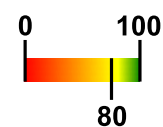
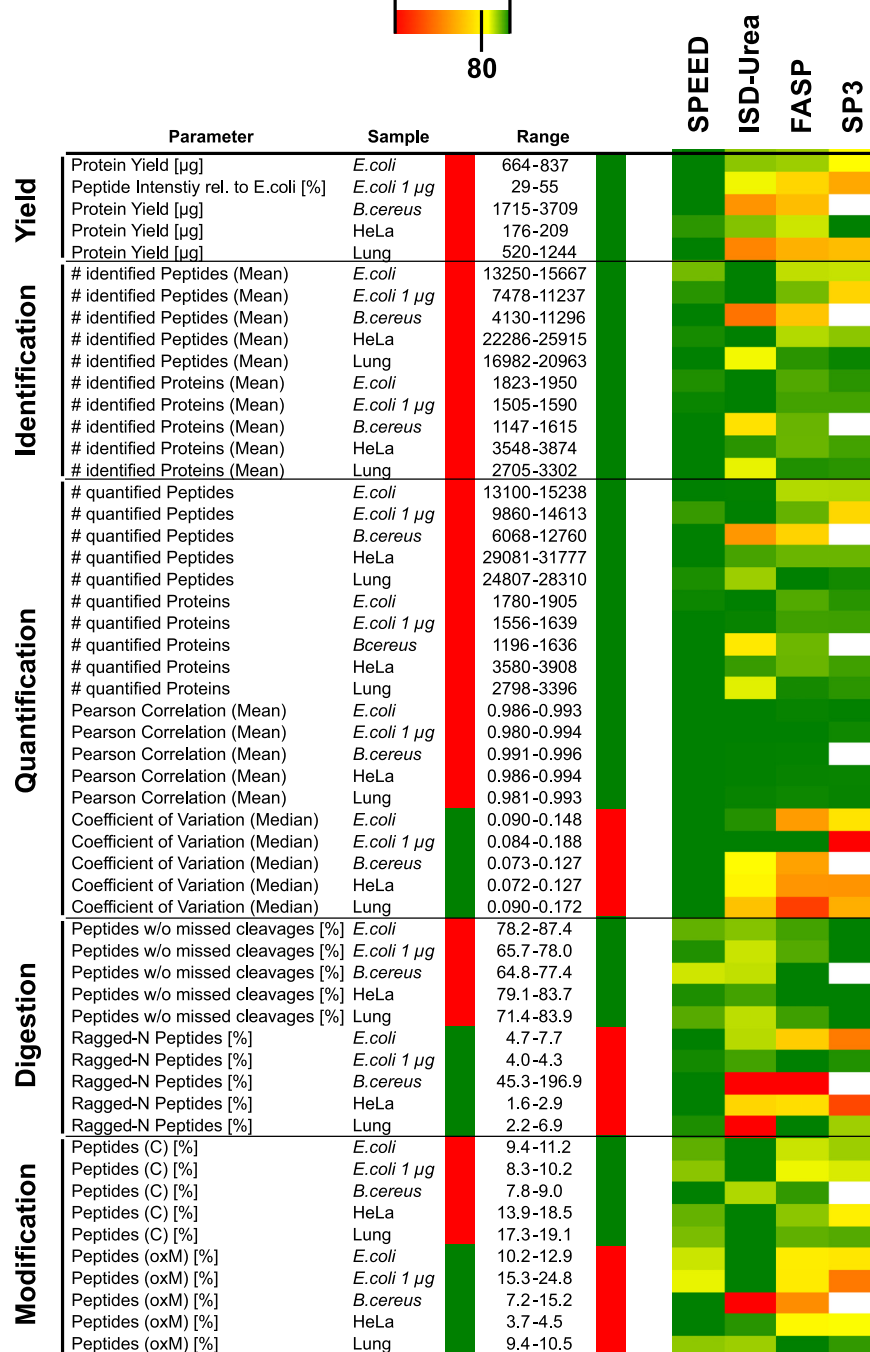


FIG. 4. Performance evaluation of SPEED with the established sample preparation methods SP3, FASP and ISD-Urea. Sample preparation by SP3, FASP, SPEED and Urea-based in-solution digestion (ISD-Urea) was compared by triplicate LFQ-based analyses of *E. coli* cells using 20 or 1 μ g starting material, *B. cereus* and HeLa cells as well as mouse lung tissue. Results of different parameters sorted according to the sections yield, identification, quantification, digestion and modifications are displayed in a heatmap. The best result for each parameter is set to 100% (dark green) and relative differences between the sample preparation methods are color-coded according to the legend. Red and green surroundings of the range covered by each parameter show if either large or small numbers are desirable. As *B. cereus* cells could not be prepared using SP3, the related sections in the heatmap are displayed white.



ing amounts (1 and 20 μ g protein), an easy-to-lyse sample of high complexity (HeLa), a difficult-to-lyse sample of high complexity (mouse lung tissue) and a highly lysis-resistant sample of medium complexity (*B. cereus*). The results of LC-MS measurements were compared by extracting 11 sample preparation dependent parameters from the data, which were categorized into the sections yield, identification, quantification, digestion and modification (Fig. 4 and [supplemental Table S1](#)).

Protein Yield—TFA-based protein extraction enables rapid and complete dissolution of cells and tissues without the use of physical disruption methods. Accordingly, SPEED increased the yield of protein extraction from difficult-to-lyse samples types by at least 68% for mouse lung tissue and at least 54% for *B. cereus* compared with detergent or chaotropic agent-based methods (Fig. 4). However, for further sample processing equal protein amounts were used for each method.

Identification—SPEED outperformed FASP and SP3 considering the number of identified proteins for all samples analyzed by 1–9% and ISD-Urea for HeLa, *B. cereus* and mouse lung tissue by 3–41% (Fig. 4 and [supplemental Table S1](#)). ISD-Urea identified most proteins in both analyses of easy-to-lyse *E. coli* cells, exceeding SPEED by 3% (20 μ g starting material) or 1% (1 μ g starting material) respectively. However, ISD-Urea identified least proteins in both more difficult to lyse sample types, mouse lung tissue (–22%) and *B. cereus* (–41%), which underlines the sample type dependence of urea for the effective extraction of proteins. Further, we were not able to prepare *B. cereus* for LC-MS analysis using SP3, most likely because the cell wall interfered with peptide generation. Exclusive protein identifications for each method were identified in HeLa cells and mouse lung tissue by enabling the transfer of identifications among all samples using the “match between run” option in MaxQuant ([supplemental Fig. S7](#)). This increases sensitivity and reduces potential biases introduced by the stochastic nature of data-dependent acquisition. However, the number of truly exclusive identifications, which could not be detected on MS¹-level in at least one replicate of any other method, was low and revealed no significantly enriched GO-terms. Instead, the observed differences between SPEED, FASP, SP3 and ISD-Urea in HeLa cells and mouse lung tissue are mainly resulting from different protein quantities ([supplemental Fig. S8](#)).

Quantification—Protein intensities correlated well for all methods with Pearson coefficients exceeding 0.98 for all sample types, but coefficients of variation (CV) varied. SPEED-based sample processing resulted in the lowest variations of protein intensities with median CVs of < 10% for all sample types. Because further sample preparation using FASP and SP3 resulted in increased CVs compared with ISD-Urea for all easy-to-lyse sample types, reproducibility of protein quantification seems to benefit from low number of sample handling steps. In accordance with this observation, SPEED quantified the largest number of proteins without missing values for all sample types except *E. coli* (20 μ g starting amount), where ISD-Urea exceeded SPEED by 1% (Fig. 4 and [supplemental Table S1](#)). Gene ontology analysis of differentially abundant proteins ([supplemental Fig. S8A–C](#)) in mouse lung tissue further revealed that SPEED is able to prepare membrane proteins for LC-MS as efficiently as SDS-based methods (SP3, FASP) in contrast to ISD-Urea ([supplemental Fig. S8D](#), [supplemental Table S2](#)).

Digestion—Tryptic digestion of proteins is the fundamental step for bottom-up proteomics sample preparation. The efficacy of proteolysis depends on the source of trypsin, protein-to-enzyme ratio, temperature, and incubation time, which were all kept constant in this study independently of the sample preparation method, as well as on protein concentration and buffer composition, which were variable (25, 26). The lowest numbers of peptides with missed-cleavage sites were detected in SP3 prepared samples in all experiments (12–

22%). Presumably this finding results from the low digestion volume compared with FASP, SPEED and ISD-Urea. However, higher digestion efficacy is not necessarily associated with lower variability in protein quantitation (27). Preparations using SPEED resulted in 2–6% higher proportions of peptides with missed-cleavage sites, but outperformed SP3 in terms of protein CVs in all experiments. The numbers of ragged-N peptides in SPEED preparations were lowest or at least second lowest (+ 0.07% relative to total number of peptide IDs) in all preparations. Again, this underlines that peptides are not hydrolyzed by acidification using pure TFA.

Peptide Modifications—Proteins are identified from mass spectra by database searching while usually considering carbamidomethylation of cysteines and oxidation of methionine as possible peptide modifications. FASP, ISD-Urea and SP3 use dithiothreitol (DTT) for reduction and iodoacetamide (IAA) for alkylation of disulfide-bonds consecutively, whereas SPEED uses TCEP and CAA simultaneously. The simultaneous use of TCEP and CAA at elevated temperatures reduces sample processing times considerably, however CAA has been found to correlate with increasing numbers of methionine oxidation, although no possible mechanism was proposed (28). For method comparison the number of observed modifications was normalized relative to the number of identified peptides in the same sample ([supplemental Table S3](#)). The largest proportion of cysteine-containing peptides was detected in ISD-Urea preparations except for *B. cereus*, however at the expense of increased over-alkylation leading to modification of undesired amino acid residues ([supplemental Fig. S9](#)). SPEED identified ~ 10% fewer carbamidomethylated cysteines from HeLa, *E. coli* and mouse lung tissue compared with Urea-ISD, but over-alkylation was observed 4–5 fold less frequently. Methionine oxidation was detected at low levels in general (maximum in all samples = 24.6%) and was further found to be independent from alkylation reagents ([supplemental Table S2](#)). Therefore, at least when reduction and alkylation reagents are solubilized immediately before use, use of TCEP and CAA had no unwanted side-effects compared with DTT and IAA.

To clarify the influence of TFA addition on peptide modifications in general, an open search analysis of the mouse lung data using wide precursor mass windows (500 Da) in MSFragger was performed (21). The resulting Peptide Spectrum Matches (PSM) per Mass bin are deposited as [supplemental Table S4](#). Modifications with more than 100 PSMs in at least one sample are summarized in [supplemental Table S5](#) along with some selected PTMs. The results of these unbiased modification analysis show, that TFA does indeed not add any distinct mass shift to peptides in a relevant amount. For instance, only between 2 and 4 PSMs of trifluoroacetylated peptides, which result from reaction of TFA with peptides, were identified from 97,785–100,983 MS² spectra the SPEED prepared samples.

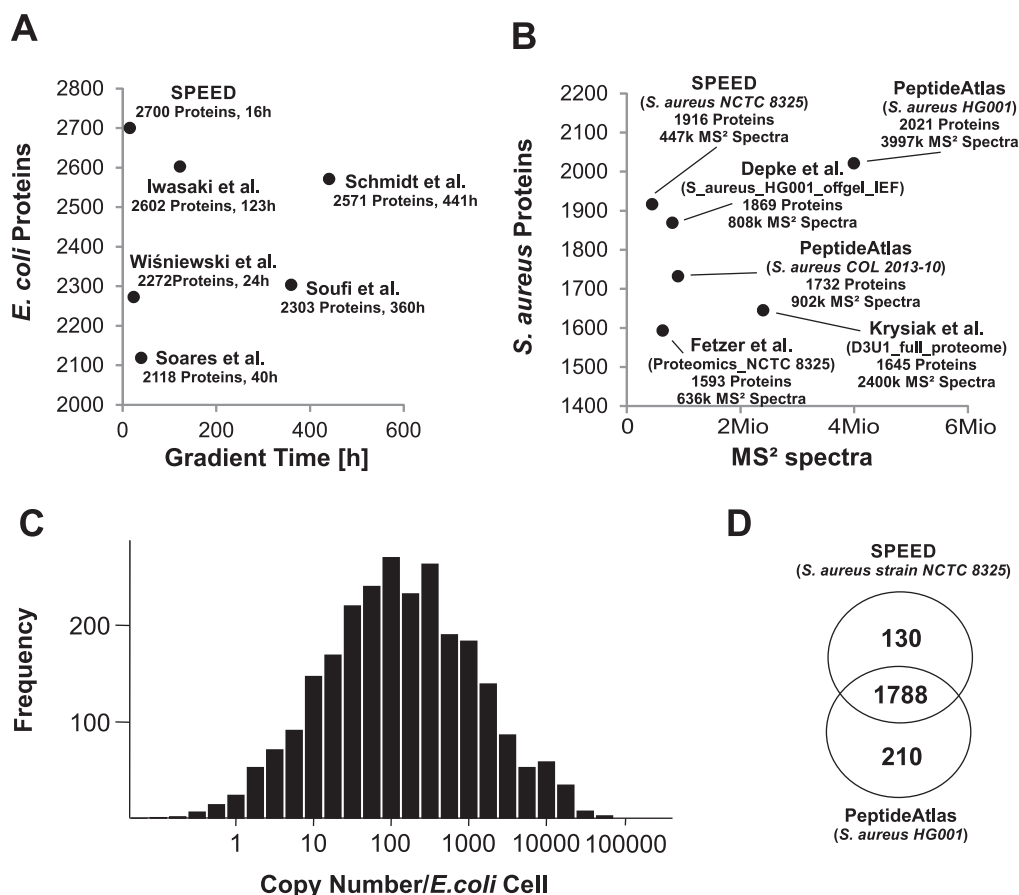


FIG. 5. Deep proteome analysis of *E. coli* and *S. aureus*. The number of protein identifications obtained from fractionated *E. coli* (A) and *S. aureus* (B) samples are compared with different large-scale and multi-sample studies including the PeptideAtlas entries of *S. aureus* (30–37). Protein copy numbers of the *E. coli* sample was calculated based on the total protein approach suggesting a lower limit of detection below 1 copy number per cell (C) (20). The overlap of the protein identifications of the SPEED-processed *S. aureus* sample and the PeptideAtlas are shown in a Venn diagram (D).

Deep Proteome Analysis of Bacteria—SPEED enables highly efficient protein extraction and peptide generation without the removal of any sample material. Therefore, it is well suited for deep and comprehensive proteome analysis. The enhancement of proteome coverage using SPEED was studied in *E. coli* and *S. aureus* cells, as the depth of analysis of bacteria depends less on the speed of the mass spectrometer and more on the sample preparation in contrast to eukaryotic cells. A single sample of either *E. coli* or *S. aureus* was fractionated using StageTips (29) and analyzed in 16 h LC-MS total gradient time. The resulting proteome coverage was compared with large-scale and multi-sample studies from the same species (Fig. 5). Using SPEED, we were able to identify more proteins from a single sample than any large-scale study of *E. coli*, even when combining the protein identifications of all samples from multiple conditions within a study (30–34). The data therefore represent the most comprehensive proteome analysis of well-studied *E. coli* cells with a detection limit below 1 copy number per cell. Further SPEED-based preparation of the Gram-positive bacteria *S. aureus* enabled us to again identify more proteins from a single-sample than

published large scale studies (35–37). The data contains only 5% less proteins than the complete PeptideAtlas (www.peptideatlas.org) entry of the same *S. aureus* strain, which was built from ~ 4 million spectra. These results illustrate the large possibilities for enhancing proteome coverage using SPEED.

Analysis of a Microbiome Standard—The microbiome is gaining increasing attention within the scientific community for being an important factor for human health (38). Diverse microbial species colonize the human body and act as chemical factories, which degrade and synthesize a plethora of metabolites and thus expand the capabilities of the human metabolism (39, 40). However, the investigation of the microbiome is a major challenge for proteomics because of its great diversity. Sample preparation is a particularly critical step in metaproteomic studies as species-specific biases can occur because of different efficacies of cell lysis and protein extraction in microbial mixtures (41). A commercial microbiome standard consisting of 20 bacterial species (~ 10 µg starting material according to turbidity measurements) was therefore selected as a test sample to demonstrate the general applicability of SPEED sample preparation and its superior per-

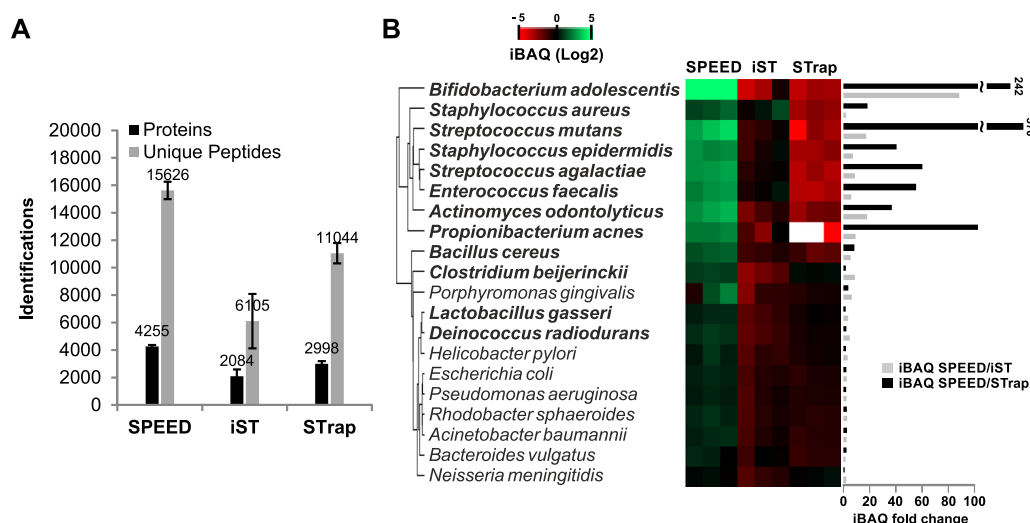


FIG. 6. Analysis of a microbiome standard. A commercial microbiome standard containing 20 bacterial species was prepared in triplicates using SPEED, STRap and iST and analyzed subsequently by 4 h single-shot LC-MS. Peptide and protein identification numbers are illustrated in (A). Unique proteins from the individual bacterial species were quantified using the summed iBAQ-intensities. Ratios of protein abundancies resulting from different sample processing protocols are visualized in a heatmap (B) in which the fold changes are color-coded. Gram-positive bacteria are marked bold.

formance compared with detergent-based methods. Because we were unable to prepare Gram-positive bacteria using SP3, ISD-Urea has proven ineffective for lysis of Gram-positive bacteria and FASP was not well suited for the preparation of low protein starting amounts, we chose to compare SPEED with the two detergent-based methods, suspension trapping (STRap) and in-StageTip (iST) (6, 10). S-Trap is based on the removal of SDS by trapping proteins in a quartz-filter after precipitation, whereas iST is an in-solution approach, which is based on the dilution of a trypsin-compatible detergent.

SPEED-based sample processing increased the number of identified unique peptide sequences by 40% (compared with STRap) and 156% (iST) and the number of identified proteins by 42% (STRap) and 104% (iST) (Fig. 6A). Absolute quantification of species-unique proteins by summed intensity-values (iBAQ) (42) revealed that especially proteins from Gram-positive bacteria were highly enriched in SPEED-samples (Fig. 6B). This indicates that TFA enhanced protein extraction from lysis-resistant bacteria up to 370-fold compared with the combination of detergents and physical disruption. *t* test analysis of the species-quantification data further showed, that 17 (compared with STRap) or 18 (iST) out of 20 bacterial species were significantly enriched ($FDR = 0.01$, $s_o = 0$) by SPEED-based sample processing. These data clearly demonstrate the superior lysis and extraction efficacy of TFA compared with detergents and the great benefit of SPEED for analysis of the microbiome.

DISCUSSION

Comprehensive, accurate and deep proteome analysis relies on the reproducible generation of peptides from the entirety of proteins present in each sample. SPEED offers sev-

eral inherent benefits over currently established methods. Because of its fundamentally different chemistry it represents a minimalistic approach for universal proteomics sample preparation. SPEED combines the straightforward nature of in-solution digestion with a highly efficient one-step extraction suitable even for membrane proteins of tissue samples or lysis-resistant Gram-positive bacteria. Adaptation of the sample preparation protocol to the respective sample type is not necessary because the mere addition of TFA is enough for efficient protein extraction, which circumvents variability problems associated with mechanical disruption and the use of detergents and chaotropic agents (43). In conjunction with online desalting, SPEED enables sample processing in a one-pot manner from cells to peptides in less than 15 min hands-on time without removal of any sample material and therefore without protein losses. Further, the detergent-free chemistry of SPEED can be combined with protein aggregation capture on spin filters, which enables sample processing in less than 40 min total time. This could be an attractive approach when time to results matter, such as pathogen identification and characterization.

The performance evaluation of SPEED for the analysis of samples with varying complexity and lysis-resistance revealed that the method outperforms detergent-based (FASP, SP3) and chaotropic agent-based (ISD-Urea) protocols in quantifying the largest number of proteins with the best accuracy for a representative selection of sample types and is also well-suited for the analysis of low protein starting amounts. The two major benefits of SPEED were found to be the enhancement of reproducibility in quantitative experiments resulting from the minimized sample handling, and the

improvement of the proteome coverage of challenging samples because of the highly efficient TFA-based protein extraction. The universal applicability of SPEED even for preparing the most challenging samples was further demonstrated by analyzing a commercial microbiome standard consisting of 20 bacterial species. SPEED was found to increase the numbers of peptide and protein identifications by > 40% and to enhance the detection of lysis-resistant Gram-positive bacteria 2–370-fold compared with detergent-based methods. Further, the coupling of SPEED with peptide fractionation enabled us to analyze the proteomes of *E. coli* and *S. aureus* with an unprecedented depth. Acidification with TFA is further known to reliably inactivate pathogenic microorganisms including highly resistant bacterial endospores from *Bacillus anthracis*, and SPEED is therefore well suited for the processing of infectious samples, including highly pathogenic agents (44).

The rapid processing, high performance and broad applicability of the SPEED protocol is realized solely on the smart use of chemicals, requires no special equipment, is inexpensive (1 €/sample) and can easily be accomplished by non-experts. Automation of sample preparation in proteomics is a prerequisite to increase throughput, enhance reproducibility and so enable routine clinical applications. SPEED possesses some unique properties for this purpose. It is the only protocol which enables lysis of different sample materials, such as tissue, eukaryotic cells or bacteria, by simply pipetting up and down without the need for physical disruption methods, which in contrast complicates automation. Further, quantification by turbidity measurement of proteins and even on-line monitoring of the digestion process is straight-forward as well and can be carried out in well-plate readers without the need for any sample manipulation, e.g. for use of colorimetric assays. The whole preparation process is based on only two procedures, the addition of liquids to the sample containing vial and the control of the sample temperature. Both procedures can theoretically be implemented even on rather low-level liquid-handling stations and should enable the automation of SPEED. Because of these inherent benefits, SPEED has the potential to be widely adopted in the proteomics community as an ultra-rapid, low-cost, detergent-free and universal sample preparation method. The protocol simplifies sample preparation, enables standardization of lysis/extraction and thus enhances reproducibility in proteomics.

DATA AVAILABILITY

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository with the dataset identifier PXD011189.

* A.S., J.D., and P.L. are the inventors of SPEED and have submitted a patent application related to SPEED. The authors declare that they have no conflicts of interest with the contents of this article.

§ This article contains [supplemental Material](#).

‡ To whom correspondence should be addressed. Tel.: +49 30–18754 2373; E-mail: Doellingerj@rki.de.

Author contributions: J.D., A.S., and P.L. designed research; J.D., A.S., and M.H. performed research; J.D. analyzed data; J.D. and P.L. wrote the paper.

REFERENCES

- Leitner, A., and Aebersold, R. (2013) SnapShot: mass spectrometry for protein and proteome analyses. *Cell* **154**, 252–252 e251
- Leon, I. R., Schwammle, V., Jensen, O. N., and Sprenger, R. R. (2013) Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. *Mol. Cell. Proteomics* **12**, 2992–3005
- Sielaff, M., Kuharev, J., Bohn, T., Hahlbrock, J., Bopp, T., Tenzer, S., and Distler, U. (2017) Evaluation of FASP, SP3, and iST protocols for proteomic sample preparation in the low microgram range. *J. Proteome Res.* **16**, 4060–4072
- Tanca, A., Biosia, G., Pagnozzi, D., Addis, M. F., and Uzzau, S. (2013) Comparison of detergent-based sample preparation workflows for LTQ-Orbitrap analysis of the *Escherichia coli* proteome. *Proteomics* **13**, 2597–2607
- Wisniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362
- Zougman, A., Selby, P. J., and Banks, R. E. (2014) Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* **14**, 1006–0. Epub 2014 Mar 26
- Duan, X., Young, R., Straubinger, R. M., Page, B., Cao, J., Wang, H., Yu, H., Canty, J. M., and Qu, J. (2009) A straightforward and highly efficient precipitation/on-pellet digestion procedure coupled with a long gradient nano-LC separation and Orbitrap mass spectrometry for label-free expression profiling of the swine heart mitochondrial proteome. *J. Proteome Res.* **8**, 2838–2850
- Hughes, C. S., Foehr, S., Garfield, D. A., Furlong, E. E., Steinmetz, L. M., and Krijgsveld, J. (2014) Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10**, 757
- Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., and Yates, J. R., 3rd. (1999) Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **17**, 676–682
- Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014) Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11**, 319–324
- Chang, Y. H., Gregorich, Z. R., Chen, A. J., Hwang, L., Guner, H., Yu, D., Zhang, J., and Ge, Y. (2015) New mass-spectrometry-compatible degradable surfactant for tissue proteomics. *J. Proteome Res.* **14**, 1587–1599
- Humphrey, S. J., Karayel, O., James, D. E., and Mann, M. (2018) High-throughput and high-sensitivity phosphoproteomics with the EasyPhos platform. *Nat. Protoc.* **13**, 1897–1916
- Masuda, T., Sugiyama, N., Tomita, M., and Ishihama, Y. (2011) Microscale phosphoproteome analysis of 10,000 cells from human cancer cell lines. *Anal. Chem.* **83**, 7698–7703
- Wang, H., Qian, W. J., Mottaz, H. M., Clauss, T. R., Anderson, D. J., Moore, R. J., Camp, D. G., 2nd, Khan, A. H., Sforza, D. M., Pallavicini, M., Smith, D. J., and Smith, R. D. (2005) Development and evaluation of a micro- and nanoscale proteomic sample preparation method. *J. Proteome Res.* **4**, 2397–2403
- Kollipara, L., and Zahedi, R. P. (2013) Protein carbamylation: in vivo modification or in vitro artefact? *Proteomics* **13**, 941–944
- Wisniewski, J. R., and Gaugaz, F. Z. (2015) Fast and sensitive total protein and Peptide assays for proteomic analysis. *Anal. Chem.* **87**, 4110–4116
- Ishihama, Y., Rappsilber, J., and Mann, M. (2006) Modular stop and go extraction tips with stacked disks for parallel and multidimensional peptide fractionation in proteomics. *J. Proteome Res.* **5**, 988–994
- Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372
- Tyanova, S., and Cox, J. (2018) Perseus: a bioinformatics platform for integrative analysis of proteomics data in cancer research. *Methods Mol. Biol.* **1711**, 133–148

20. Wisniewski, J. R., and Mann, M. (2016) A proteomics approach to the protein normalization problem: selection of unvarying proteins for MS-based proteomics and Western blotting. *J. Proteome Res.* **15**, 2321–2326
21. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D., and Nesvizhskii, A. I. (2017) MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* **14**, 513–520
22. Katz, J. J. (1954) Anhydrous trifluoroacetic acid as a solvent for proteins. *Nature* **174**, 509
23. Wang, A. Y., Thuy-Boun, P. S., Stupp, G. S., Su, A. I., and Wolan, D. W. (2019) Triflic acid treatment enables LC-MS/MS analysis of insoluble bacterial biomass. *J. Proteome Res.* **17**, 2978–2986
24. Batth, T. S., Tollenaere, M. A. X., Ruther, P., Gonzalez-Franquesa, A., Prabhakar, B. S., Bekker-Jensen, S., Deshmukh, A. S., and Olsen, J. V. (2019) Protein aggregation capture on microparticles enables multipurpose proteomics sample preparation. *Mol. Cell Proteomics* **18**, 1027–1035
25. Betancourt, L. H., Sanchez, A., Pla, I., Kuras, M., Zhou, Q., Andersson, R., and Marko-Varga, G. (2018) Quantitative assessment of urea in-solution Lys-C/trypsin digestions reveals superior performance at room temperature over traditional proteolysis at 37 degrees C. *J. Proteome Res.*
26. Hildonen, S., Halvorsen, T. G., and Reubsæet, L. (2014) Why less is more when generating tryptic peptides in bottom-up proteomics. *Proteomics* **14**, 2031–2041
27. Chiva, C., Ortega, M., and Sabido, E. (2014) Influence of the digestion technique, protease, and missed cleavage peptides in protein quantitation. *J. Proteome Res.* **13**, 3979–3986
28. Hains, P. G., and Robinson, P. J. (2017) The Impact of Commonly Used Alkylating Agents on Artifactual Peptide Modification. *J. Proteome Res.* **16**, 3443–3447
29. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906
30. Iwasaki, M., Miwa, S., Ikegami, T., Tomita, M., Tanaka, N., and Ishihama, Y. (2010) One-dimensional capillary liquid chromatographic separation coupled with tandem mass spectrometry unveils the Escherichia coli proteome on a microarray scale. *Anal. Chem.* **82**, 2616–2620
31. Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrne, E., Volkmer, B., Callipo, L., Knoop, K., Bauer, M., Aebersold, R., and Heinemann, M. (2016) The quantitative and condition-dependent Escherichia coli proteome. *Nat. Biotechnol.* **34**, 104–110
32. Soares, N. C., Spat, P., Krug, K., and Macek, B. (2013) Global dynamics of the Escherichia coli proteome and phosphoproteome during growth in minimal medium. *J. Proteome Res.* **12**, 2611–2621
33. Soufi, B., Krug, K., Harst, A., and Macek, B. (2015) Characterization of the E. coli proteome and its modifications during growth and ethanol stress. *Front. Microbiol.* **6**, 103
34. Wisniewski, J. R., and Rakus, D. (2014) Quantitative analysis of the Escherichia coli proteome. *Data Brief* **1**, 7–11
35. Depke, M., Michalik, S., Rabe, A., Surmann, K., Brinkmann, L., Jehmlich, N., Bernhardt, J., Hecker, M., Wollscheid, B., Sun, Z., Moritz, R. L., Volker, U., and Schmidt, F. (2015) A peptide resource for the analysis of Staphylococcus aureus in host-pathogen interaction studies. *Proteomics* **15**, 3648–3661
36. Fetzer, C., Korotkov, V. S., Thanert, R., Lee, K. M., Neuenschwander, M., von Kries, J. P., Medina, E., and Sieber, S. A. (2017) A chemical disruptor of the ClpX chaperone complex attenuates the virulence of multidrug-resistant Staphylococcus aureus. *Angew Chem. Int. Ed. Engl.* **56**, 15746–15750
37. Krysiak, J., Stahl, M., Vomacka, J., Fetzer, C., Lakemeyer, M., Fux, A., and Sieber, S. A. (2017) Quantitative map of beta-lactone-induced virulence regulation. *J. Proteome Res.* **16**, 1180–1192
38. Gilbert, J. A., Blaser, M. J., Caporaso, J. G., Jansson, J. K., Lynch, S. V., and Knight, R. (2018) Current understanding of the human microbiome. *Nat. Med.* **24**, 392–400
39. Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., and Gordon, J. I. (2007) The human microbiome project. *Nature* **449**, 804–810
40. Locey, K. J., and Lennon, J. T. (2016) Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5970–5975
41. Blackburn, J. M., and Martens, L. (2016) The challenge of metaproteomic analysis in human samples. *Expert Rev. Proteomics* **13**, 135–138
42. Schwanhauss, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature* **473**, 337–342
43. 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., Smith, R. D., and Myers, A. J. (2013) Sources of technical variability in quantitative LC-MS proteomics: human brain tissue sample analysis. *J. Proteome Res.* **12**, 2128–2137
44. Lasch, P., Nattermann, H., Erhard, M., Stämmler, M., Grunow, R., Bannert, N., Appel, B., and Naumann, D. (2008) MALDI-TOF mass spectrometry compatible inactivation method for highly pathogenic microbial cells and spores. *Anal. Chem.* **80**, 2026–2034