A Protein Epitope Signature Tag (PrEST) Library Allows SILAC-based Absolute Quantification and Multiplexed Determination of Protein Copy Numbers in Cell Lines*[S]

Marlis Zeiler‡, Werner L. Straube‡**, Emma Lundberg§, Mathias Uhlen§¶, and Matthias Mann‡‖**

Mass spectrometry-based proteomics increasingly relies on relative or absolute quantification. In relative quantification, stable isotope based methods often allow mixing at early stages of sample preparation, whereas for absolute quantification this has generally required recombinant expression of full length, labeled protein standards. Here we make use of a very large library of Protein Epitope Signature Tags (PrESTs) that has been developed in the course of the Human Protein Atlas Project. These PrESTs are expressed recombinantly in *E. coli* and they consist of a short and unique region of the protein of interest as well as purification and solubility tags. We first quantify a highly purified, stable isotope labeling of amino acids in cell culture (SILAC)-labeled version of the solubility tag and use it determine the precise amount of each PrEST by its SILAC ratios. The PrESTs are then spiked into cell lysates and the SILAC ratios of PrEST peptides to peptides from endogenous target proteins yield their cellular quantities. The procedure can readily be multiplexed, as we demonstrate by simultaneously determining the copy number of 40 proteins in HeLa cells. Among the proteins analyzed, the cytoskeletal protein vimentin was found to be most abundant with 20 million copies per cell, while the transcription factor and oncogene FOS only had 6000 copies. Direct quantification of the absolute amount of single proteins is possible via a SILAC experiment in which labeled cell lysate is mixed both with the heavy labeled solubility tag and with the corresponding PrEST. The SILAC-PrEST combination allows accurate and streamlined quantification of the absolute or relative amount of proteins of interest in a wide variety of applications. *Molecular & Cellular Proteomics 11: 10.1074/mcp.O111.009613, 1–13, 2012.*

MS-based proteomics has become a method of choice to study proteins in a global manner (1–3). Mass spectrometry is not inherently quantitative but many methods have been developed to overcome this limitation. Most of them are based on stable isotopes and introduce a mass shifted version of the peptides of interest, which are then quantified by their "heavy" to "light" ratio. Stable isotope labeling is either accomplished by chemical addition of labeled reagents, enzymatic isotope labeling, or metabolic labeling (4–6). Generally, these approaches are used to obtain relative quantitative information on proteome expression levels in a light and a heavy labeled sample. For example, stable isotope labeling by amino acids in cell culture (SILAC)[1] (7, 8) is performed by metabolic incorporation of light or heavy labeled amino acids into the proteome. Labeled proteomes can also be used as internal standards for determining protein levels of a cell or tissue proteome of interest, such as in the spike-in SILAC approach (9).

Absolute quantification is technically more challenging than relative quantification and can only be performed accurately for a single or a small number of proteins at a time (10). Typical applications of absolute quantifications are the determination of cellular copy numbers of proteins (important for systems biology) or the concentration of biomarkers in body fluids (important for medical applications). Furthermore, any precise method of absolute quantification, when performed in more than one sample, also yields the relative amounts of the protein between these samples.

Several methods for absolute quantification have emerged over the last years including absolute quantification (AQUA) (11), quantification concatamer (QConCAT) (12, 13), protein standard absolute quantification (PSAQ) (14), absolute SILAC (15), and FlexiQuant (16). They all quantify the endogenous protein of interest by the heavy to light ratios to a defined

From the ‡Department of Proteomics and Signal Transduction, Max-Planck Institute of Biochemistry, D-82152 Martinsried, Germany; §Science for Life Laboratory, Royal Institute of Technology, SE-17165 Stockholm, Sweden; ¶AlbaNova University Center, Royal Institute of Technology, SE-10691 Stockholm, Sweden

⌘ *Author's Choice*—Final version full access.

Received March 15, 2011, and in revised form, September 22, 2011

Published, MCP Papers in Press, September 30, 2011, DOI 10.1074/mcp.O111.009613

[1] The abbreviations used are: SILAC, stable isotope labeling of amino acids in cell culture; AQUA, absolute quantification; PSAQ, protein standard absolute quantification; QconCAT, quantification concatamer; PrEST, protein epitope signature tag; AIF, all ion fragmentation; ABP, albumin binding protein.

amount of the labeled counterpart spiked into the sample and are chiefly distinguished by either spiking in heavy labeled peptides or heavy labeled full length proteins. The AQUA strategy is convenient and streamlined: proteotypic peptides (17) are chemically synthesized with heavy isotopes and spiked in after sample preparation. AQUA peptides are commercially available but currently relatively expensive, especially when many peptides or proteins need to be quantified. More fundamentally, the AQUA strategy suffers from quantification uncertainties that are introduced because of spiking in of the peptide standard after sample preparation and enzymatic proteolysis, which is a late stage in the workflow. Furthermore, any losses of the peptides—for example during storage—would directly influence quantification results. The QconCAT approach is based on artificial proteins that are concatamers of proteotypic peptides. This artificial protein is recombinantly expressed in *Escherichia coli* and spiked into the sample before proteolysis. QconCAT in principle allows efficient production of labeled peptides but does not automatically correct for protein fractionation effects or digestion efficiency in the native proteins *versus* the concatamers. The PSAQ, absolute SILAC and FlexiQuant approaches sidestep these limitations by metabolically labeling full length proteins by heavy versions of the amino acids arginine and lysine. PSAQ and FlexiQuant *in vitro* synthesize full-length proteins in wheat germ extracts or in bacterial cell extract, respectively, whereas absolute SILAC was described with recombinant protein expression in *E. coli*. The protein standard is added at an early stage, such as directly to cell lysate. Consequently, sample fractionation can be performed in parallel and the SILAC protein is digested together with the proteome under investigation. However, these advantages come at the cost of having to produce full length proteins, which limits throughput and generally restricts these methods to soluble proteins.

In this study we advance the absolute SILAC approach by making use of a highly scalable and already established system for protein standard production. We employ short Protein Epitope Signature Tags (PrESTs), which are produced in a high-throughput manner by the Human Protein Atlas project and subsequently used as antigens for antibody production (18–20). The ultimate goal of the Human Protein Atlas is to produce at least two specific antibodies to all human proteins and to use this resource to study the tissue distribution and the subcellular distribution of the human proteome (21–24). PrESTs incorporate a sequence of about 100 amino acids of the target protein chosen for minimal homology to other proteins. Other criteria include avoidance of signal peptide sequences and sequences from transmembrane spanning regions. These PrEST sequences are fused to a 6xHis tag for purification and to a solubility tag derived from the albumin binding domain of the streptococcal protein G (25). We reasoned that these attributes, combined with the fact that more than 30,000 PrESTs, representing 18,300 human genes have already been produced, would make the PrEST library an excellent resource for streamlined, absolute SILAC based quantification of human proteins.

EXPERIMENTAL PROCEDURES

Protein Epitope Signature Tags—The short protein fragments were produced in high-throughput by the Human Protein Atlas where they are used as antigens for antibody production (26, 27). In brief, suitable Protein Epitope Signature Tags (PrESTs) representing unique regions of each target protein were designed using the human genome sequence as template (EnsEMBL). Unique PrESTs with a size between 50 to 150 amino acids and low homology to other human proteins were selected, excluding epitope- and domain-sized similarities to other proteins, signal peptides, and transmembrane regions (26). The cloning, protein expression, and purification were performed as previously described (27, 28). Part of the quality control is that all PrESTs are evaluated and purity verified using SDS-PAGE and the molecular weight is determined by mass spectrometry before further use. This also excludes major "laddering" of the PrESTs. For optimal storage PrESTs were lyophilized and dissolved in 8 M urea and stored at −20 °C until further use. To ascertain that the PrESTs had an endogenous counterpart in HeLa cells, we selected 50 proteins spread over the abundance range of a HeLa proteome that we had measured at a depth of about 4000 proteins. Proteins were picked without regards to specific protein classes, cellular localizations, or functions. Of these 50 proteins, 43 were readily available from the Protein Atlas pipeline in recombinantly expressed form. For multiplexing experiments these 43 PrESTs were mixed together—each at the appropriate concentration. This "master mix" that was then spiked into cell lysates.

Cell Culture—For SILAC labeling, HeLa cells were cultured in Dulbecco's modified Eagle's medium (Invitrogen, Carlsbad, CA) containing 10% dialyzed fetal bovine serum (Invitrogen) and penicillin/streptomycin (Invitrogen). Heavy arginine (high purity Arg10, Cambridge Isotope Laboratories, Andover, MA) and heavy lysine (high purity Lys8, Cambridge Isotope Laboratory) were added to a final concentration of 33 μg/ml or 76 μg/ml, respectively. After six passages cells were fully labeled as assessed by mass spectrometry. Cells were counted using a Countess cell counter (Invitrogen) and aliquots of 10^6 cells were snap frozen and stored at −80 °C.

Protein Expression and Purification of ABP (Albumin Binding Protein)—The expression vector pAff8c (Human Protein Atlas) was modified via SLIC cloning (29) inserting a OneStrep affinity tag to the C terminus of the Albumin Binding Protein (ABP). To express heavy labeled ABP in *E. coli*, an expression strain auxotrophic for arginine and lysine was used (40). Cultures were grown in PA5052 minimal autoinduction media as previously described in (30) but with the addition of 18 normal (light) amino acids and heavy arginine and lysine. Cultures were grown overnight and harvested at an OD600 of about 5.7. *E. coli* cells were lysed in 100 mM Tris, 150 mM NaCl and Protease Inhibitor (Roche) using a Bioruptor (Diagenode, Denville, NY). Cell debris was removed by centrifugation and the supernatant was cleared by filtration through a 22-μm filter. The soluble ABP was purified by affinity chromatography on a StrepTap Hitrap column (GE Healthcare) coupled to an ÄKTA system. The protein was loaded in binding buffer (100 mM Tris, 150 mM NaCl, 1 mM dithiothreitol). After washing with 10 column volumes it was eluted with elution buffer (100 mM Tris, 150 mM NaCl, 1 mM dithiothreitol, 2.5 mM desthiobiotin) (31). The purity of the protein was evaluated by mass spectrometry via an in-solution digest followed by liquid chromatography tandem MS (LC MS/MS). Abundances of ABP and contaminants were estimated by adding the signal for their most intense peptides. ABP was dialyzed in phosphate-buffered saline (PBS), aliquoted, snap-frozen and stored at −80 °C. The concentration of purified ABP was measured by amino acid analysis (Genaxxon BioScience GmbH).

The heavy labeled PrESTs were produced as described above except that the PrEST was expressed and labeled using the auxotrophic *E. coli* strain in media containing Arg10 and Lys8.

Sample Preparation—HeLa cells were lysed in 100 mM Tris, 4% SDS, 100 mM dithiotreitol, incubated for 5 min at 95 °C and disrupted using a Bioruptor. The lysate was cleared by centrifugation through SpinX filters (22 μm, Corning, Corning, NY). The PrESTs were added at appropriate concentrations (see main text) to labeled HeLa cells and the samples were further processed by the FASP method (32). In brief, proteins were captured on a 30-kDa filter and SDS was exchanged with an urea containing buffer. Proteins were alkylated with iodoacetamide and trypsinized (Promega, Charbonnières, France). Further peptide separation was performed using pipette-based six fraction SAX as described (33).

The PrESTs and ABP were mixed and solubilized in denaturation buffer (6 M urea, 2 M thiourea in 10 mM HEPES, pH 8), reduced with dithiotreitol and subsequently alkylated with iodoacetamide. The protein mixture was digested with LysC (Wako) for 3 h, diluted with ammonium bicarbonate and further digested with trypsin overnight. The digestion was stopped by acidifying with trifluoroacetic acid and desalted on C_{18}-Empore disc StageTips (34).

Liquid Chromatography and Mass Spectrometry—Analysis of the PrESTs spiked into HeLa cells was performed on an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) coupled to an Easy nano-HPLC via a nanoelectrospray ion source (Proxeon Biosystems, now Thermo Fisher Scientific). The peptides were separated on a 20 cm fused silica emitter packed in-house with reversed phase material ReproSil-Pur 120 C18-AQ 1.8 μm resin (Dr. Maisch GmbH) and eluted with a 205-min gradient from 5–35% buffer B (80% acetonitrile, 0.5% acetic acid). The mass spectrometer was operated in a data dependent fashion to automatically measure MS and consecutive MS/MS. LTQ-Orbitrap full scan MS spectra (from 100 or 300 to 1650 *m/z*) were acquired with a resolution of 60,000 at *m/z* 400. The ten most abundant ions were sequentially isolated and fragmented using higher energy collisional dissociation (HCD) followed by analysis in the Orbitrap (35).

The ratios of the light PrEST *versus* heavy ABP peptides were analyzed online on the Exactive instrument with HCD option (Thermo Fisher Scientific) using the same nano-HPLC setup as described above. The peptides were eluted with a linear gradient with 5–30% buffer B over 40 min. The Exactive mass spectrometer identified peptides with all ion fragmentation (AIF) by performing alternating MS scans (300–1600 *m/z*) of the precursor ions and all ion fragmentation scans (100–1600 *m/z*) using stepped HCD fragmentation (36). Both scans were acquired at a resolution of 100,000 at *m/z* 200.

The heavy PrESTs *versus* light ABP peptides were analyzed using the TriVersa Nanomate (Advion Biosciences), a chip implementation of nanoelectrospray coupled to a LTQ Orbitrap XL. The samples were eluted in 50% methanol, 0.5% formic acid. A voltage of 1.6 kV and a nitrogen gas pressure of 0.35 psi was applied to spray the peptides into the mass spectrometer. Each sample is sprayed to a single nozzle on the electrospray ionization chip eliminating carryover. A standard data dependent top10 collision-induced dissociation fragmentation method was applied for 2 min acquiring ~40 full scans for quantification (37).

Data Analysis—Acquired data were analyzed with MaxQuant (38) (version 1.2.0.11) using the human IPI database (version 3.68; 87,083 entries). Common contaminants and the sequence of the ABP solubility tag were added to this database. For peptide identification we used Andromeda, a probabilistic search engine incorporated in to the MaxQuant framework (39). Carbamidomethylation of cysteine was included in the search as a fixed modification and methionine oxidation as well as N-terminal acetylation were included as variable modifications. We allowed two miscleavages and required a minimum of six amino acids per identified peptide. The initial mass tolerance for

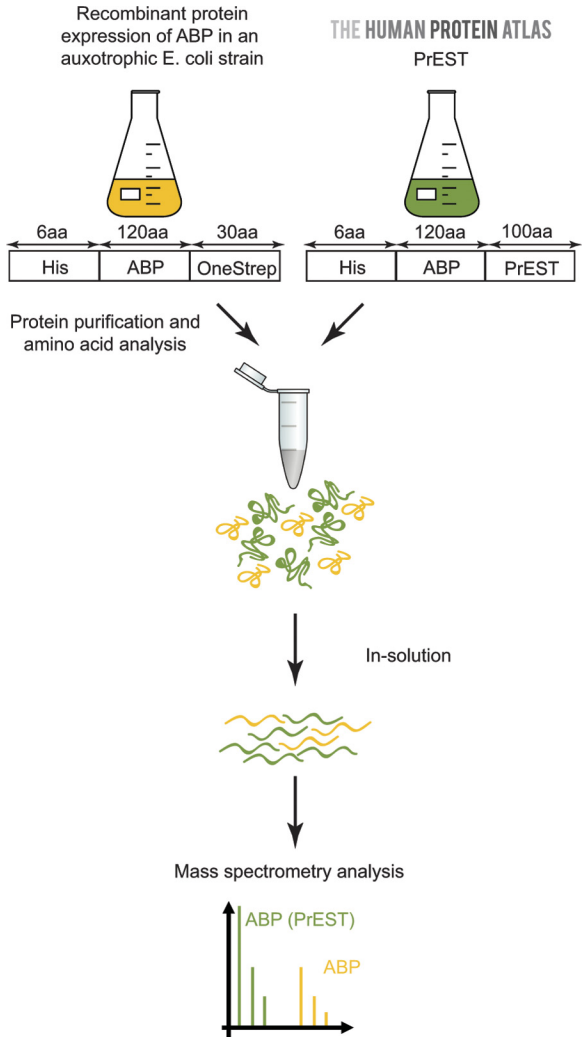

FIG. 1. **Schematic workflow for accurate determination of PrEST concentrations.** Heavy or light ABP is recombinantly expressed in an auxotrophic *E. coli* strain and purified using the C-terminal OneStrep tag. The heavy labeled ABP, whose concentration is measured separately by amino acid analysis, and the PrEST are mixed together and an in-solution digest is performed. Peptides are measured with a short LC MS/MS run on a benchtop mass spectrometer and the PrEST concentration is accurately determined by the SILAC ratio of the ABP peptides originating from the PrEST and the ABP.

precursor ions or fragment ions was set to 6 ppm and fragment masses were allowed to deviate by up to 0.5 Th. For statistical evaluation of the data obtained, the posterior error probability and false discovery rate (FDR) were used. The false discovery rate was determined by searching a reverse database and was set to 0.01 for peptide identification. Additional peptides were identified by the "match between run" option in MaxQuant, which matches precursor masses in a 2-min retention time window (after realignment of the runs) based on the accurate mass.

The AIF data was processed as described above except that up to 50 peaks were analyzed per 100 *m/z* with a tolerance of 15 ppm. The precursor ion mass was matched with the possible fragment ion candidates on the basis of the cosine correlation value of at least 0.6 (36).

Enzyme-linked Immunosorbent Assay—Absolute amount measurements of proto-oncogene c-Fos and Stratifin (14–3-3 σ) was carried

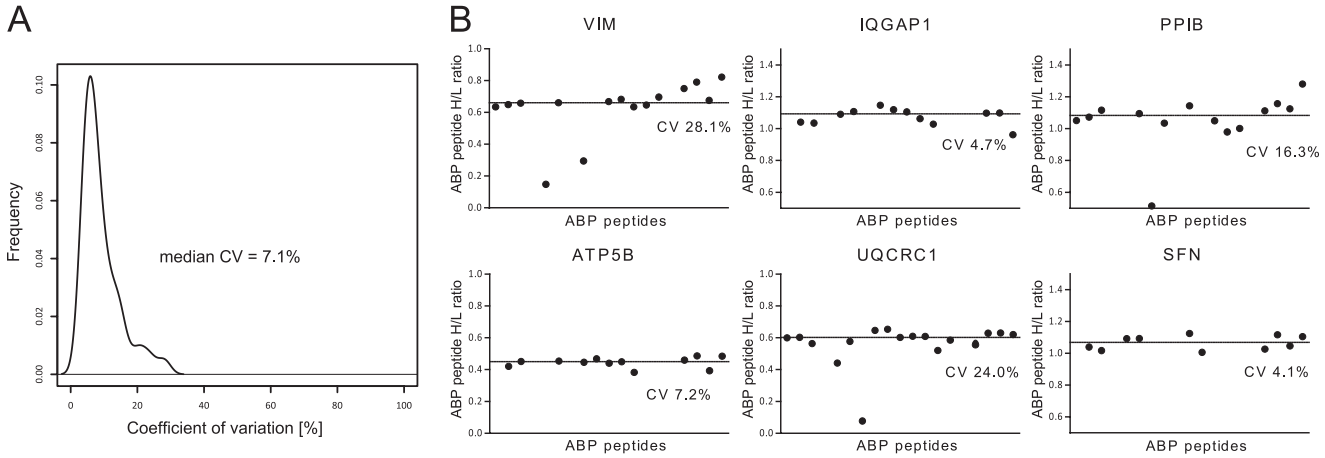

FIG. 2. **Accuracy of ABP quantification.** *A,* Density plot of the overall distribution of the 43 coefficients of variation (CVs) of the ABP peptides measured on a benchtop Exactive mass spectrometer. *B*, Representative example proteins showing the H/L peptide ratios of the ABP peptides deriving from the ABP standard and the ABP peptides in the PrESTs and their coefficients of variation (CVs).

out by ELISA. The kits were purchased from USCNK Life Science and performed according to the manufacturer's instructions. The HeLa cells were lysed in PBS, RIPA 1 (50 mM Tris pH 7.5, 150 mM NaCl, 1% Nonidet P-40) or RIPA 2 (50 mM Tris pH 7.5, 150 mM NaCl, 1% Nonidet P-40; 0.1% SDS) with protease inhibitors. The cells were disrupted by 3 freeze-thaw cycles and sonication using the Biorupter. For the ELISA the samples were diluted 1:10. Fluorescence activity was measured by a microplate reader (Tecan) and converted to actual concentration by a standard curve.

RESULTS

Unlike relative quantification, absolute quantification is a two-step process that requires measurement of firstly the absolute amount of the standard and secondly the relative amount of the standard compared with the analyte of interest. Determination and subsequent control of the level of standard is by no means trivial and can easily be the step that limits the overall accuracy of the approach. Below, we first describe a generic method to determine the absolute amount of each PrEST with high accuracy. Then we construct a "master mix" of different PrESTs and evaluate the ability of the SILAC - PrEST method to accurately quantify cellular proteins. We then apply the master mix to determine the copy numbers of 40 proteins in HeLa cells, a human cervical carcinoma cell line. Finally, we describe an alternative workflow for the quantification of single proteins of interest, in which the two steps are combined into one LC MS/MS analysis.

Accurate Measurement of PrEST Concentrations—Each PrEST is already fused to the ABP, a solubilization tag of 120 amino acids. *In silico* digest of ABP results in 40 tryptic peptides with a length between 6 and 30 amino acids (supplemental Table S1). We recombinantly expressed a heavy SILAC labeled version of the ABP protein tag. When necessary, we used a dual affinity approach based on an N-terminal His-tag and a C-terminal OneStrep tag to generate highly purified protein fragment and to ensure that only full length ABP was obtained. The absolute concentration of ABP protein fragment was determined by amino acid analysis, which is the most accurate method for protein quantification, but which is only applicable to highly purified proteins in relatively large amounts. Heavy SILAC incorporation into ABP was 99% and its purity was about 97% as judged by mass spectrometry (see Experimental Procedures). Because these two factors operate in a compensating direction and because of the small size of the effect, the measured concentration of ABP was not adjusted for them.

LC MS/MS of ABP indeed revealed many readily detectable tryptic peptides (see below). Each of the 43 PrESTs from the Protein Atlas Project was separately mixed with a known amount of labeled ABP as schematically outlined in Fig. 1 to allow for a SILAC LC-MS/MS experiment. As this experiment requires a separate LC MS/MS run for each PrEST it was likely to be rate-limiting for the overall project. We therefore decided to perform this analysis on an economical and robust benchtop Orbitrap instrument rather than on a Velos instrument. The Exactive instrument cannot isolate peptide precursors, therefore we identified the peptides by AIF (36) in 1-h runs.

Typically, at least eight labeled ABP peptides could be quantified against the corresponding ABP peptides from the PrESTs, leading to a median coefficient of variation (CV) of 7% for PrEST quantification (Fig. 2*A*).

To overcome the step of measuring the PrESTs concentration, which limits overall throughput, the heavy PrESTs were measured by static nanoelectrospray on an automated chip-based system (TriVersa Nanomate). This enabled higher throughput measurements of these simple mixtures of ABP peptides using low sample amounts. The peptide ratio showed a median coefficient of variation 5.5%, an improvement over the Exactive based measurement of 7%.

Importantly, a particular PrEST quantification can be repeated at this stage until a desired accuracy is achieved. Here, this was not done, because the accuracy of PrEST

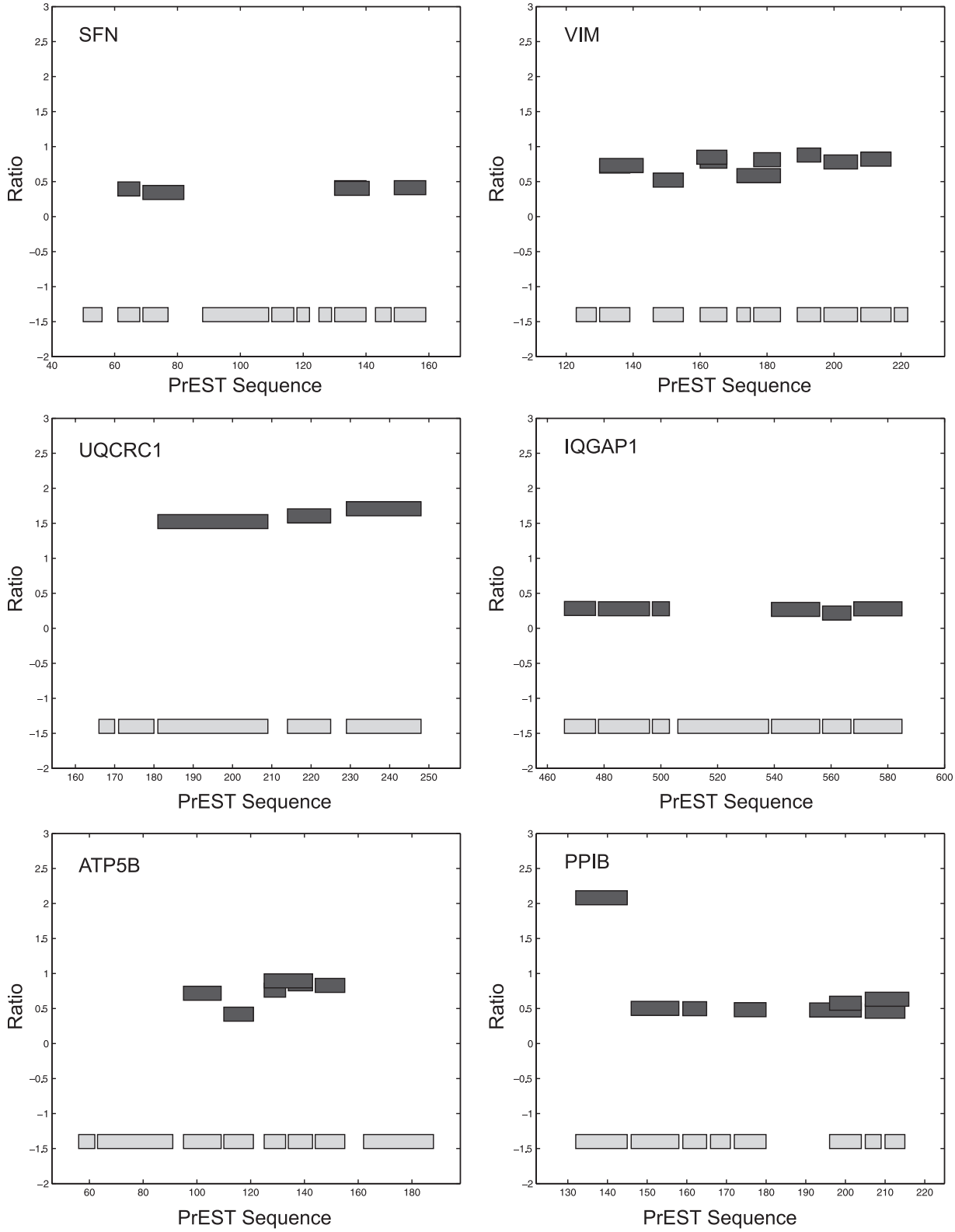

FIG. 3. **Peptide ratios along the PrESTs sequences.** The PrEST master mix was spiked into a lysate of HeLa cells and measured against the endogenous protein. The peptide ratios were extracted to quantify the proteins. The variation of the peptide ratios along the sequence is depicted. Overlapping peptides are because of missed cleavages. The gray bars correspond to the predicted limit tryptic peptides for the PrEST region.

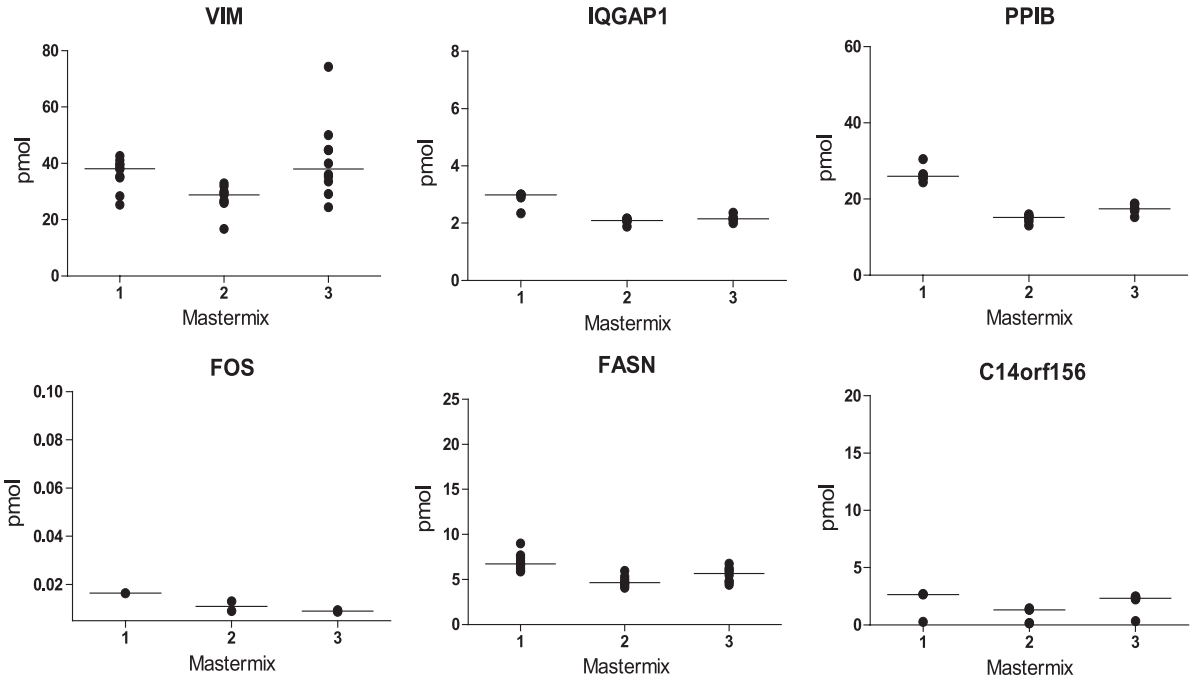

FIG. 4. **Reproducibility of the absolute quantification procedure.** Three independent quantification experiments for representative examples, in which the master mix preparation as well as the PrEST quantification were performed independently. The bars reflect the median of the peptide ratios for each protein.

quantification was estimated to be higher than that of the other steps in the workflow. A few typical examples of results from the PrEST quantification are shown in Fig. 2*B*. The median of the SILAC ABP ratios is used for robustness. This largely eliminates the contribution of outliers such as the ABP peptide (ISEATDGLSDFLKSQTPAEDTVK) in the PrESTs for proteins PPIB and UQCRC1, which had signals very close to noise levels. Note that the quantification accuracy does not depend on the cellular abundance or any other attributes of the target protein, because the same amounts of PrEST is used in each PrEST quantification experiment. Importantly, quantification accuracy in our workflow also does not depend on the purity of the PrEST because our method specifically measures the concentration of PrEST and not of total protein.

PrEST Master Mix and Endogenous Protein Quantification—Having quantified the PrEST amounts we proceeded to measuring protein expression levels in HeLa cells. For convenience we first used unlabeled PrESTs and quantified against heavy SILAC labeled HeLa cells. Because digested total cell lysates consist of hundreds of thousands of tryptic peptides, the addition of a single or even a large number of PrESTs does not change the overall complexity of the mixture. On the basis of the quantitative amounts established above, we here mixed 43 PrESTs together. In initial experiments we used equimolar mixtures of PrESTs, which were spiked into HeLa lysate in different amounts. The measured SILAC ratios established appropriate levels of each PrEST in the master mix, such that the SILAC ratios were within the most accurately quantifiable range, *i.e.* relatively close to one to one.

The master mix with appropriate levels of all the 43 PrESTs was spiked into the lysate of SILAC labeled cells. The mixture was digested according to the FASP protocol followed by SAX fractionation and resulting in six fractions that were separately measured with 4-h gradients on an LTQ Orbitrap Velos mass spectrometer. We were able to quantify 40 of the 43 proteins targeted by our PrEST master mix. Proteins were generally quantified with several PrEST derived peptides (average 3.98 and median 3), leading to an overall median CV of 12% (supplemental Table S2). As an example, the adhesion protein IQGAP1 was quantified with six peptides, which each gave nearly identical quantification results (CV 9.9%). Five of the six quantified tryptic peptides of ATP5B (mitochondrial ATP synthase subunit beta), had very close SILAC ratios, however, one peptide had a ratio that differed by 43% from the median. This peptide is clearly an outlier and its deviating value contributes substantially to the CV value, raising it from 8.7% to 23%. Note however, that we base protein quantification on the median of the peptide values; therefore the outlier peptide hardly contributes to the measured protein expression value and the CV value therefore underestimates the accuracy actually obtained in this experiment. For the same reason modifications of the endogenous proteins in the region covered by the PrEST could cause outlier peptide ratios, which would contribute little to the measured protein ratio (Fig. 3).

To independently assess the precision of this step of absolute protein quantification, we compared the ratios determined from "limit tryptic peptides" (those without internal arginine or lysine) to those determined from the longer ver-

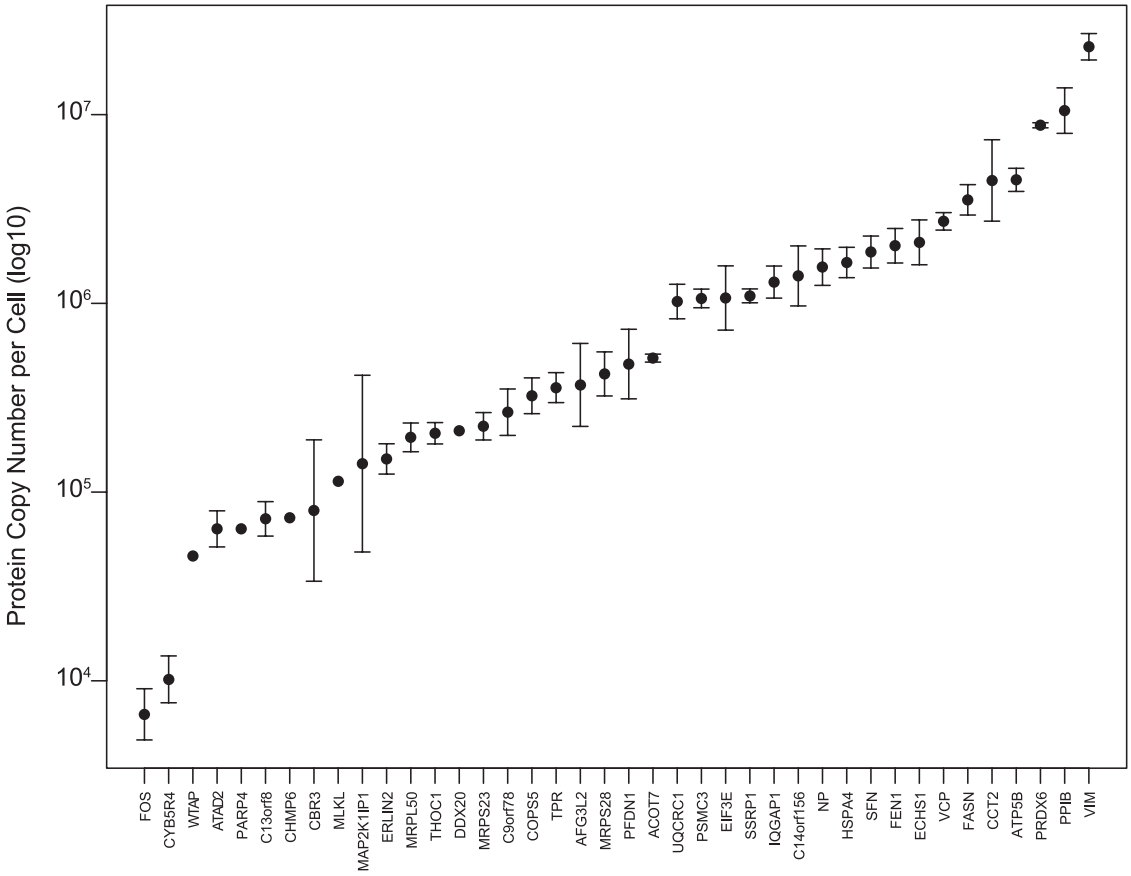

FIG. 5. **Protein copy numbers determined per HeLa cell.** The dot plot shows the protein copy numbers per cell measured in three independent experiments. The error bars correspond to the CVs. Proteins with copy numbers ranging from 6000 to 20,000,000 per cell were quantified (see also Table II).

sions of the peptide containing one or two missed tryptic cleavage sites. These peptides are very problematic for peptide standard based methods such as AQUA, but in our measurements very similar ratios were measured for such peptides. This shows that digestion proceeded identical for PrEST and endogenous protein (Table I). Thus, far from introducing uncertainty, in the SILAC-PrEST approach these peptides can provide additional quantification information.

To assess the degree of variability associated with both steps of the absolute quantification procedure, we repeated the entire workflow two more times, including PrEST quantification and master mix generation as well as measurement of cellular abundance of the target proteins. This analysis showed that the standard errors of the mean associated with all steps together are on average 20%. This value is excellent and to our knowledge the most accurate determination of cellular expression levels reported so far. Even more importantly, the errors of each of the steps in the workflow for each of the proteins are immediately apparent from the individual CVs. Thus all protein expression level measurements can be classified and accepted or discarded according to the confidence of measurements. Fig. 4 displays typical examples of protein expression determination from the triplicate measurements. Comparing the peptide ratio spreads to the variability of the mean protein values revealed that the preparation of the master mix contributed the largest variability whereas errors because of SILAC ratio determination were somewhat lower. Automated preparation of the master mix could therefore lead to further improvements in the future.

Protein Copy Number Determination in HeLa Cells—Next we used the absolute values for protein amounts in our HeLa cell lysate to calculate the corresponding copy numbers in cells. HeLa cells numbers were determined automatically in a cell counter (see Experimental Procedures). Given the known amount of each PrEST and their SILAC ratios with respect to the endogenous proteins we determined the cellular copy numbers of 40 different proteins. Very high accuracy of absolute quantification to within a standard error of 25% was achieved for 35 of 40 proteins (Table II).

Cellular copy numbers are only known for very few proteins and it is therefore interesting to relate these copy numbers to the known functions of the proteins (supplemental Table S3). The cytoskeletal protein vimentin forms intermediate filaments and was the most abundant protein with 20 million copies per cell. At the other extreme, the transcription factor and oncogene FOS is present in about 6000 copies in our HeLa cell

TABLE I

Comparison of limit tryptic peptides and peptides with missed tryptic cleavage sites. Peptides with one or two miscleavages as well as their ratios are depicted. The ratios of the two versions vary on average by 12%, which is in the normal range of variation of peptides derived from one protein

Gene name	Sequence	Missed cleavages	Ratio H/L mastermix (1)	CV (%)	Ratio H/L mastermix (2)	CV (%)	Ratio H/L mastermix (3)	CV (%)
ATP5B	IPVGPETLGR	0	0.85299	8.13	0.84263	16.92	1.1617	8.76
ATP5B	VLDSGAPIK	0	0.76219		0.67515		0.99543	
ATP5B	VLDSGAPIKIPVGPETLGR	1	0.89528		0.95159		1.1652	
CCT2	ILIANTGMDTDK	0	0.47498	39.26	0.37474	48.58	1.0965	–
CCT2	ILIANTGMDTDKIK	1	0.26858		0.1831		–	
CCT2	VAEIEHAEK	0	0.4707	6.37	0.45578	8.24	1.1598	10.43
CCT2	VAEIEHAEKEK	1	0.51512		0.51219		1.3444	
ECHS1	KLFYSTFATDDR	1	–	6.52	0.13372	23.26	–	0.01
ECHS1	LFYSTFATDDR	0	0.16792		0.15756		1.1803	
ECHS1	LFYSTFATDDRK	1	0.18416		0.20966		1.1801	
FASN	QQEQQVPILEK	0	0.73946	4.04	0.69517	2.02	0.8985	1.63
FASN	RQQEQQVPILEK	1	0.69837		0.67562		0.87805	
FEN1	LDPNKYPVPENWLHK	1	0.73377	1.04	0.68	2.58	1.2048	3.81
FEN1	YPVPENWLHK	0	0.72303		0.70528		1.1416	
HSPA4	EDQYDHLDAADMTK	0	0.2685	14.36	0.20351	5.45	0.79843	3.91
HSPA4	NKEDQYDHLDAADMTK	1	0.21899		0.21981		0.84382	
PPIB	DKPLKDVIIADCGK	2	0.47962	12.85	0.41764	14.10	1.0739	14.23
PPIB	DVIIADCGK	0	0.5755		0.51014		1.3142	
PRDX6	ELAILLGMLDPAEK	0	0.77082	14.35	0.61275	4.80	1.2449	1.05
PRDX6	ELAILLGMLDPAEKDEK	1	0.62879		0.65579		1.2636	
PRDX6	VVFVFGPDK	0	0.6815	12.56	0.74577	2.37	1.1822	15.71
PRDX6	VVFVFGPDKK	1	0.81434		0.72122		0.94586	
SFN	YLAEVATGDDK	0	0.41281	1.53	0.38665	5.75	0.98489	15.17
SFN	YLAEVATGDDKK	1	0.404		0.41942		0.79403	
TPR	LESALTELEQLR	0	0.1666	3.45	0.15943	8.07	1.1901	8.15
TPR	LESALTELEQLRK	1	0.17493		0.17872		1.3357	
VCP	DHFEEAMR	0	0.20879	1.21	0.20611	10.27	1.4096	2.44
VCP	RDHFEEAMR	1	0.20524		0.17819		1.459	
VCP	KYEMFAQTLQQSR	1	0.12676	33.99	NaN	–	0.47634	72.12
VCP	YEMFAQTLQQSR	0	0.20698		0.24338		1.4679	
VIM	QVDQLTNDK	0	0.79102	4.92	0.72609	2.57	1.3024	3.52
VIM	RQVDQLTNDK	1	0.84806		0.70013		1.2391	
VIM	EKLQEEMLQR	1	NaN	–	0.77968	7.15	NaN	–
VIM	LQEEMLQR	0	0.87912		0.8627		1.8464	
VIM	ILLAELEQLK	0	0.7214	0.83	0.71178	–	1.3289	49.04
VIM	ILLAELEQLKGQGK	1	0.7299		NaN		2.7399	
VIM	DNLAEDIMR	0	0.813	23.04	0.79552	1.43	1.6492	29.83
VIM	VEVERDNLAEDIMR	1	0.58525		0.77955		1.0746	

sample. As expected, proteins involved in cell signaling are generally expressed at lower values—as an example even the scaffolding factor mitogen-activated protein kinase scaffold protein 1 (MAP2K1IP1) is present at only 140,000 copies. However, ubiquitous signaling factors with a general chaperone-like role—such as 14-3-3 isoforms—are very highly expressed (14-3-3 sigma; 1.8 million copies). Two members of the mitochondrial ribosome have about 200,000 copies in this cell line (S23 and S28), whereas a third (L50) has about 400,000 (Note that not all ribosomal protein subunits have equal stoichiometry). The mitochondrial genome only encodes 13 genes therefore it is perhaps surprising that proteins involved in their translation are needed in such high copy numbers. A member of the respiratory chain, ATP5B, has about 4.5 million copies per HeLa cells—about fivefold higher than PSMC3, a regulatory component of the proteasome. The T-complex is a member of a chaperone system and as expected it has a very high copy number (about 4.5 million). Fatty acid synthase, a classical enzyme, is expressed at 3.5 million copies, whereas another enzyme acyl coenzyme A thioester hydrolase (ACOT7) is expressed about sevenfold lower (500,000 copies). Such expression numbers could be interesting for modeling metabolic pathways. These are anecdotal examples but they illustrate that knowledge of the absolute expression levels of cellular proteins can contribute to the understanding of their roles in the cell.

Absolute Quantification Using Heavy PrESTs—Above we used already expressed and purified PrESTs and quantified

TABLE II
Protein copy numbers per HeLa cell

Protein names	Gene name	Median	RSD (%)[a]	Mastermix 1	Mastermix 2	Mastermix 3
14-3-3 protein sigma	SFN	1,870,568	19.81	2,364,005	1,870,568	1,604,145
26S protease regulatory subunit 6A	PSMC3	1,062,048	11.37	1,062,048	950,200	1,192,875
28S ribosomal protein S23, mitochondrial	MRPS23	223,198	17.26	223,198	203,672	282,020
28S ribosomal protein S35, mitochondrial	MRPS28	422,825	24.80	473,409	284,783	422,825
39S ribosomal protein L50, mitochondrial	MRPL50	194,935	18.14	177,937	250,001	194,935
AFG3-like protein 2	AFG3L2	369,737	41.68	369,737	412,509	165,983
ATP synthase subunit beta, mitochondrial	ATP5B	4,511,967	14.68	5,672,473	4,376,424	4,511,967
ATPase family AAA domain-containing protein 2	ATAD2	63,835	23.40	63,835	61,373	91,846
Carbonyl reductase [NADPH] 3	CBR3	79,823	94.26	79,823	61,399	322,454
Charged multivesicular body protein 6	CHMP6	83,028	67.19	122,476	43,581	-
Coiled-coil domain-containing protein 55	CCDC55	–[b]	–	–	–	–
COP9 signalosome complex subunit 5	COPS5	323,791	22.62	323,791	284,218	435,937
Cytochrome b5 reductase 4	CYB5R4	10,180	30.80	16,205	10,180	9,515
Cytochrome b-c1 complex subunit 1, mitochondrial	UQCRC1	1,022,450	19.50	1,022,450	713,318	1,025,854
Cytosolic acyl coenzyme A thioester hydrolase	ACOT7	512,746	4.79	512,746	472,208	514,556
Endoplasmic reticulum lipid raft-associated protein 2	ERLIN2	149,867	19.53	206,262	148,785	149,867
Enoyl-CoA hydratase, mitochondrial	ECHS1	2,105,336	28.10	2,965,394	1,723,133	2,105,336
Eukaryotic translation initiation factor 3 subunit 6	EIF3E	1,067,627	34.63	1,067,627	599,306	1,253,469
FACT complex subunit SSRP1	SSRP1	1,095,695	8.52	1,095,695	1,022,209	1,209,724
Fatty acid synthase	FASN	3,536,145	17.98	4,043,129	2,804,853	3,536,145
Flap endonuclease 1	FEN1	2,019,699	20.42	2,372,346	2,019,699	1,563,785
Heat shock 70 kDa protein 4	HSPA4	1,646,549	19.22	2,146,713	1,499,858	1,646,549
Hepatocellular carcinoma-associated antigen 59	C9orf78	265,003	25.76	289,516	171,397	265,003
Lysophosphatidylcholine acyltransferase 1	AYTL2	–	–	–	–	–
Mitogen-activated protein kinase scaffold protein 1	MAP2K1IP1	141,520	68.85	182,796	27,116	141,520
Mixed lineage kinase domain-like protein	MLKL	114,801	17.14	128,711	–	100,891
Nucleoprotein TPR	TPR	357,637	17.53	397,408	278,736	357,637
Peptidyl-prolyl cis-trans isomerase B	PPIB	10,502,199	29.14	15,610,836	9,112,850	10,502,199
Peroxiredoxin 6	PRDX6	8,781,079	3.07	8,881,373	8,377,838	8,781,079
Poly [ADP-ribose] polymerase 4	PARP4	63,971	7.07	60,775	67,168	–
Prefoldin subunit 1	PFDN1	476,849	36.22	476,849	523,643	243,332
Pre-mRNA-splicing regulator WTAP	WTAP	49,143	51.10	31,385	–	66,902
Probable ATP-dependent RNA helicase DDX20	DDX20	213,466	19.17	242,403	184,529	–
Proto-oncogene c-Fos	FOS	6,643	32.41	9,956	6,643	5,359
Purine nucleoside phosphorylase	NP	1,555,814	23.04	2,101,680	1,357,920	1,555,814
Ras GTPase-activating-like protein IQGAP1	IQGAP1	1,296,511	20.65	1,796,903	1,260,937	1,296,511
SRA stem-loop-interacting RNA-binding protein, mitochondrial	C14orf156	1,397,500	32.95	1,665,787	828,707	1,397,500
T-complex protein 1 subunit beta	CCT2	4,479,130	48.47	7,447,762	2,757,533	4,479,130
THO complex subunit 1	THOC1	204,962	13.16	239,173	184,576	204,962
Transitional endoplasmic reticulum ATPase	VCP	2,719,254	10.44	2,719,254	2,358,278	2,904,468
Uncharacterized protein C1orf65	C1orf65	–	–	–	–	–
Vimentin	VIM	22,886,339	15.22	22,974,646	17,376,010	22,886,339
Zinc finger protein 828	C13orf8	72,135	19.47	74,281	51,084	72,135

[a] Standard error of the mean (S.E.) for the three replicates in percent.
[b] No valid data obtained.

against heavy ABP protein and heavy SILAC-labeled cell lysate. Although convenient to determine copy numbers in cell lines, in other applications it would be more appropriate to express heavy labeled PrESTs, which can then be mixed into any proteome of choice—including tissue and clinical body fluid samples. To apply our absolute quantification approach to nonlabeled samples we expressed 28 of the PrESTs in heavy SILAC labeled *E. coli*, purified them and prepared a heavy master mix. To streamline quantification of PrEST levels, we developed an automated set up employing static nanoelectrospray (Advion NanoMate; see EXPERIMENTAL PROCEDURES). As expected, spiking the heavy master mix into normal, non-SILAC labeled cells allowed equally straightforward quantification of the targeted proteins, with good correlation to the previous experiment (Fig. 6).

Absolute Quantification in Single Experiments—We also wished to develop a variation on the SILAC-PrEST strategy to quantify single protein target. In this case, the two experimental steps involved in absolute protein quantification can be collapsed into one as outlined schematically in Fig. 7*A*. A

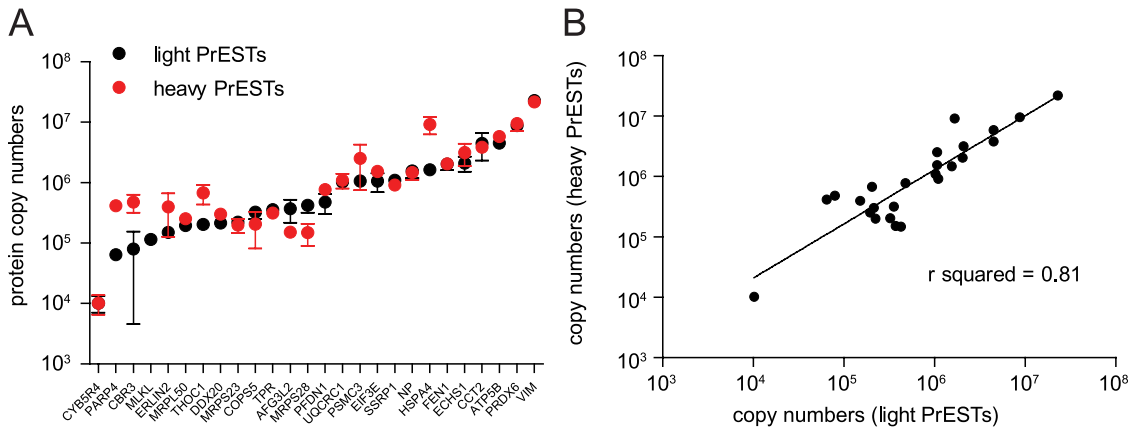

FIG. 6. **Absolute Quantification using heavy PrESTs.** *A*, Comparison of copy numbers obtained by quantifying light PrESTs against SILAC labeled heavy cell lysate (black symbols) *versus* quantifying heavy PrESTs against unlabeled cell lysate (red symbols). *B*, Values shown in *A* but plotted as a scatter graph.

precisely known amount of the ABP solubility tag is mixed into cell lysate together with the labeled PrEST. LC-MS/MS analysis of the sample then provides SILAC ratios of light ABP solubility tag to labeled PrEST ABP peptides. These ratios accurately quantify the amount of PrEST that was used. The same LC MS data also contain the ratios of labeled PrEST peptides to the unlabeled endogenous protein counterpart. Together, these ratios quantify the absolute amount of endogenous protein in a single experiment (Fig. 7*B*). Note that triple-SILAC labeling is not required in this approach because the ratios are determined against different regions of the PrEST construct, namely the common ABP solubility tag region (for quantifying the PrEST) and the protein specific PrEST region (for quantifying the endogenous protein).

This single-plex method for quantification was performed for three different HeLa proteins in which the SILAC-labeled cell lysate and SILAC-labeled ABP was quantified against unlabeled PrESTs. As shown in Fig. 7*C*, consistent values were obtained in these measurements based on triplicate experiments. The absolute levels generally agreed well with the copy numbers determined independently in the multiplexed PrEST-SILAC experiment described above (maximum difference between the means of 40%), validating both approaches.

Enzyme-linked Immunosorbent Assay—ELISA is a standard method in biochemical research to determine absolute amounts, or at least to reproducibly determine protein levels. We therefore compared the SILAC-PrEST method to this established technology. When performing the ELISA assay for Stratifin (14-3-3 σ) under typical conditions—filtered cell lysate and PBS as recommended by the manufacturer—the ELISA recorded less than 20% of the amount quantified by MS. (Note that there is no interference by 14-3-3 isoforms because these peptides are different.) The recommendation of the manufacturer to dissolve in PBS could not solubilize the pellet. The solubility was increased by adding the nonionic detergent Nonidet P-40, which was able to dissolve most of the sample pellet. Adding a low concentration of sodium dodecyl sulfate (SDS), an anionic detergent, further improvement significantly increased measured protein amount (Fig. 8*A*). Still the absolute amounts were underestimated twofold compared with mass spectrometry analysis, presumably because the FASP protocol enables complete solubilization by the use of 4% SDS.

We also investigated the levels of the transcription factor and proto-oncogene FOS by ELISA, the lowest abundance protein quantified in our mix. Here solubilization did not appear to be an issue and we received excellent agreement between quantitative values determined by MS and by ELISA using different buffer conditions (Fig. 8*B*).

DISCUSSION AND OUTLOOK

Here we have developed methods to determine the absolute levels of proteins in cells by taking advantage of the absolute SILAC concept as well as the availability of a large library of PrEST protein fragments. PrESTs already contain an ABP solubility tag, which facilitates recombinant expression of PrESTs against a wide variety of cellular targets. We found that this solubility tag is an excellent "quantification tag" because it generates a large number of readily quantifiable peptides upon tryptic digestion. After producing a highly purified and accurately quantified "gold standard" of the ABP tag alone, it can be used to quantify all PrESTs in turn. Importantly, the purity of the PrEST is not a concern because quantification is only performed against the PrEST component and not against possible *E. coli* or other contaminants.

PrESTs have already been produced against 80% of the human proteome and an "industrialized pipeline" for their production is in place. Although almost all PrESTs so far have been produced for human target proteins, they could in principle be made for any other species in exactly the same way. Furthermore, in many cases more than one PrEST has been made for the same protein to allow for the generation of paired antibodies and pair-wise validation of antibody staining patterns (23). Like-

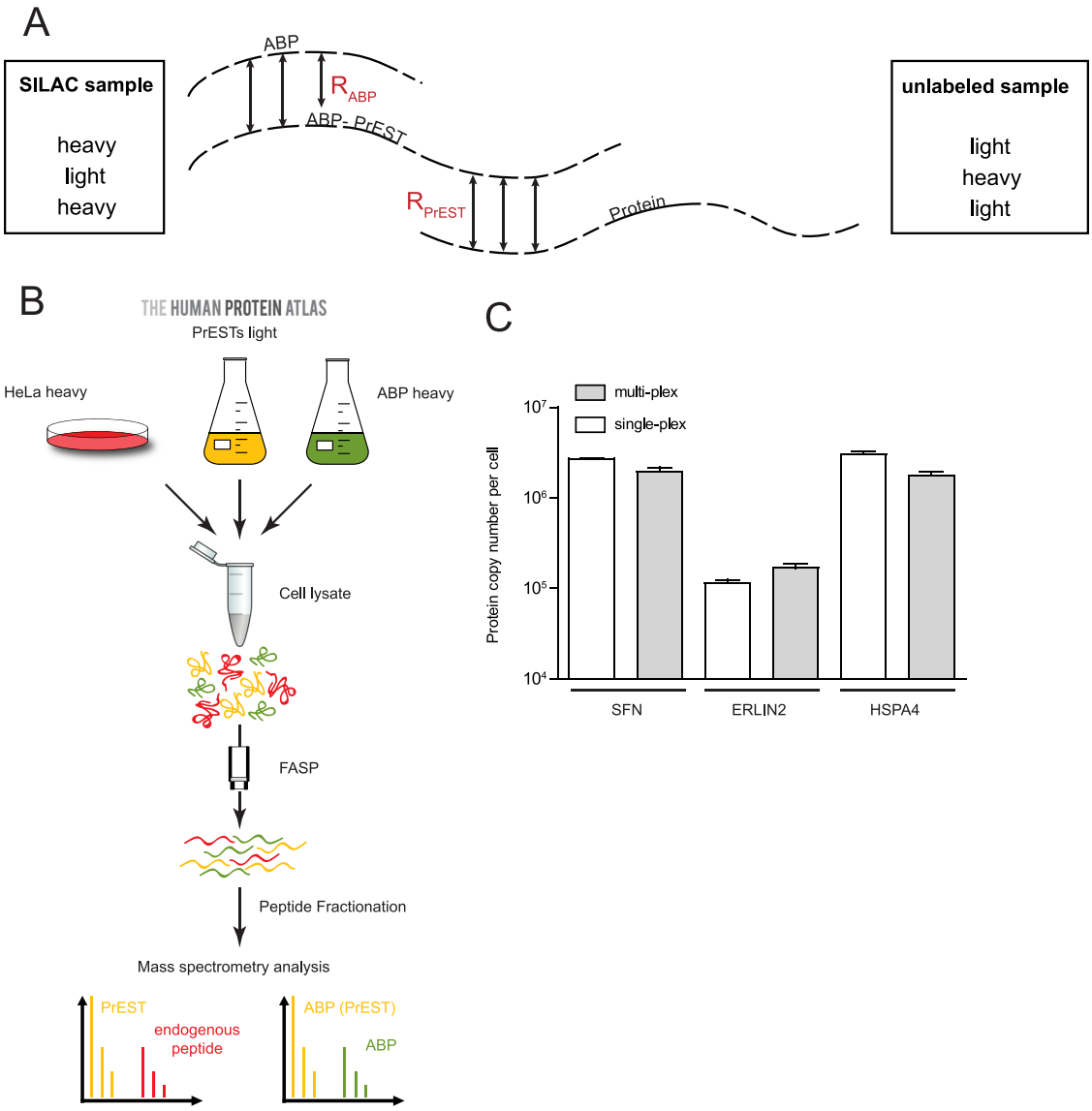

FIG. 7. **Direct quantification of a single protein in HeLa cell lysate.** *A*, Principle of the 'single-plex' strategy for the direct quantification of a single protein. In the same experiment, SILAC peptide ratios mapping to the ABP quantification tag determine the amount of PrEST whereas SILAC ratios mapping to the protein specific region of the PrEST construct determine the level of the endogenous proteins. The experiment can be performed with SILAC heavy labeled cells, unlabeled PrEST construct and heavy labeled ABP tag (*left* side) or *vice versa* (*right* side). *B*, Single-plex determination of absolute protein amount. In the workflow depicted here, an unlabeled PrEST construct as well as a heavy labeled ABP tag are both spiked into HeLa cell lysate before digestion. *C*, Comparison of copy numbers obtained from the "master mix" experiment with those from the single-plex experiments for three different proteins. Error bars are standard deviations of the mean from triplicate measurements.

wise, different PrESTs could be produced in cases where the current ones are not optimal for MS-based quantification.

For preparation of a SILAC-PrEST mixture we estimated the appropriate amount of PrESTs in a two-step procedure. In the future it may be simpler to estimate the amount of protein roughly from the peptide signals from in-depth proteome experiments. PrESTs spiked in at corresponding amounts will likely be in the easily quantifiable range in most cases. The overall accuracy of the SILAC-PrEST approach can be monitored for each step in the procedure and it currently appears to be limited by manual pipetting accuracy. It is likely that precision of all steps in the procedure can be improved significantly in the future.

Here we have demonstrated applications of SILAC-PrESTs for copy number determination in cell lines. However, the principle should be applicable in the same way to absolutely quantify proteins from any source and we plan to investigate this shortly.

Acknowledgments—We thank Shubin Ren for support with bioinformatic analysis, Christoph Schaab for providing Fig. 3, Sabine Suppmann for technical support and Tami Geiger for helpful discussions.

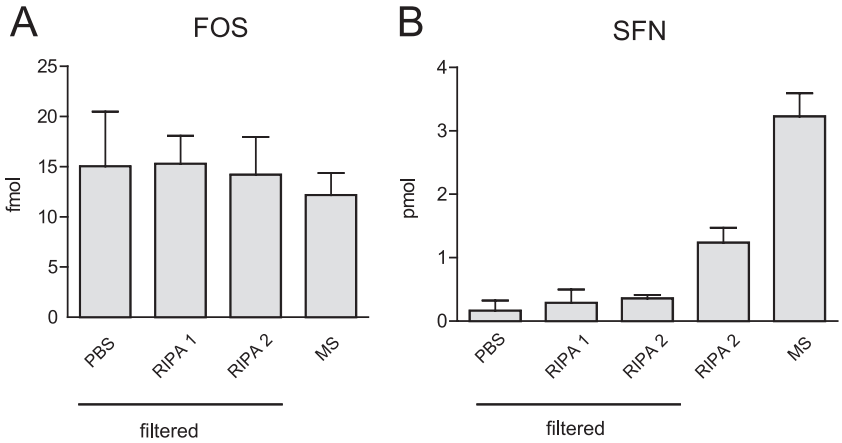

FIG. 8. **Comparison of SILAC-PrEST based quantification and ELISA.** Proto-oncogene c-Fos (*A*) and Stratifin (*B*) were quantified by ELISA to evaluate the SILAC-PrEST absolute quantification. Different ELISA compatible buffers and filtered *versus* unfiltered cell lysates were compared.

The *E. coli* strain auxotrophic for arginine and lysine was a kind gift from Ron Hay, University of Dundee.

Data availability - The acquired raw data was uploaded to Tranche (https://proteomecommons.org/tranche/) as "PrEST-SILAC absolute quantification" and is available via following hash code: HI8Dw7I jHuMUVLLXhNxbD/sJPWqKUuxUVCqa7ZncEXr/f30ksWIJPzy0/PKZ 2EyTgzRP6dU5E3BeeKy+1sUOI2ad0AAAAAAACP6Q==.

* This work was supported by the European Commission's 7th Framework Program PROSPECTS (HEALTH-F4-2008–201648, the Munich Center for Integrated Protein Science (CIPSM) and the Knut and Alice Wallenberg Foundation.

** Both authors made equal contributions to this work.

[S] This article contains supplemental Tables S1 to S4.

‖ To whom correspondence should be addressed: Department of Proteomics and Signal Transduction, Max-Planck Institute of Biochemistry, D-82152 Martinsried, Germany. Tel.: + 49-89-8578-2557; E-mail: mmann@biochem.mpg.de.

REFERENCES

1. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422,** 198–207
2. Cravatt, B. F., Simon, G. M., and Yates, J. R., 3rd (2007) The biological impact of mass-spectrometry-based proteomics. *Nature* **450,** 991–1000
3. Gstaiger, M., and Aebersold, R. (2009) Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nat. Rev. Genet.* **10,** 617–627
4. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* **1,** 252–262
5. Bachi, A., and Bonaldi, T. (2008) Quantitative proteomics as a new piece of the systems biology puzzle. *J. Proteomics* **71,** 357–367
6. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* **389,** 1017–1031
7. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1,** 376–386
8. Mann, M. (2006) Functional and quantitative proteomics using SILAC. *Nat. Rev. Mol. Cell Biol.* **7,** 952–958
9. Geiger, T., Wisniewski, J. R., Cox, J., Zanivan, S., Kruger, M., Ishihama, Y., and Mann, M. (2011) Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. *Nat. Protoc.* **6,** 147–157
10. Brun, V., Masselon, C., Garin, J., and Dupuis, A. (2009) Isotope dilution strategies for absolute quantitative proteomics. *J. Proteomics* **72,** 740–749
11. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 6940–6945
12. Beynon, R. J., Doherty, M. K., Pratt, J. M., and Gaskell, S. J. (2005) Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nat. Methods* **2,** 587–589
13. Pratt, J. M., Simpson, D. M., Doherty, M. K., Rivers, J., Gaskell, S. J., and Beynon, R. J. (2006) Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nat. Protoc.* **1,** 1029–1043
14. Brun, V., Dupuis, A., Adrait, A., Marcellin, M., Thomas, D., Court, M., Vandenesch, F., and Garin, J. (2007) Isotope-labeled protein standards: toward absolute quantitative proteomics. *Mol. Cell. Proteomics* **6,** 2139–2149
15. Hanke, S., Besir, H., Oesterhelt, D., and Mann, M. (2008) Absolute SILAC for accurate quantitation of proteins in complex mixtures down to the attomole level. *J. Proteome Res.* **7,** 1118–1130
16. Singh, S., Springer, M., Steen, J., Kirschner, M. W., and Steen, H. (2009) FLEXIQuant: a novel tool for the absolute quantification of proteins, and the simultaneous identification and quantification of potentially modified peptides. *J. Proteome Res.* **8,** 2201–2210
17. Kuster, B., Schirle, M., Mallick, P., and Aebersold, R. (2005) Scoring proteomes with proteotypic peptide probes. *Nat. Rev. Mol. Cell. Biol.* **6,** 577–583
18. Uhlén, M., Björling, E., Agaton, C., Szigyarto, C. A., Amini, B., Andersen, E., Andersson, A. C., Angelidou, P., Asplund, A., Asplund, C., Berglund, L., Bergström, K., Brumer, H., Cerjan, D., Ekström, M., Elobeid, A., Eriksson, C., Fagerberg, L., Falk, R., Fall, J., Forsberg, M., Björklund, M. G., Gumbel, K., Halimi, A., Hallin, I., Hamsten, C., Hansson, M., Hedhammar, M., Hercules, G., Kampf, C., Larsson, K., Lindskog, M., Lodewyckx, W., Lund, J., Lundeberg, J., Magnusson, K., Malm, E., Nilsson, P., Odling, J., Oksvold, P., Olsson, I., Oster, E., Ottosson, J., Paavilainen, L., Persson, A., Rimini, R., Rockberg, J., Runeson, M., Sivertsson, A., Sköllermo, A., Steen, J., Stenvall, M., Sterky, F., Strömberg, S., Sundberg, M., Tegel, H., Tourle, S., Wahlund, E., Waldén, A., Wan, J., Wernérus, H., Westberg, J., Wester, K., Wrethagen, U., Xu, L. L., Hober, S., and Pontén, F. (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteomics* **4,** 1920–1932
19. Nilsson, P., Paavilainen, L., Larsson, K., Odling, J., Sundberg, M., Andersson, A. C., Kampf, C., Persson, A., Al-Khalili Szigyarto, C., Ottosson, J., Bjorling, E., Hober, S., Wernérus, H., Wester, K., Pontén, F., and Uhlen, M. (2005) Towards a human proteome atlas: high-throughput generation of mono-specific antibodies for tissue profiling. *Proteomics* **5,** 4327–4337
20. Berglund, L., Björling, E., Oksvold, P., Fagerberg, L., Asplund, A., Szigyarto, C. A., Persson, A., Ottosson, J., Wernérus, H., Nilsson, P., Lundberg, E., Sivertsson, A., Navani, S., Wester, K., Kampf, C., Hober, S., Pontén, F., and Uhlén, M. (2008) A genecentric Human Protein Atlas for expression

profiles based on antibodies. *Mol. Cell. Proteomics* **7,** 2019–2027
21. Pontén, F., Gry, M., Fagerberg, L., Lundberg, E., Asplund, A., Berglund, L., Oksvold, P., Björling, E., Hober, S., Kampf, C., Navani, S., Nilsson, P., Ottosson, J., Persson, A., Wernérus, H., Wester, K., and Uhlén, M. (2009) A global view of protein expression in human cells, tissues, and organs. *Mol. Syst. Biol.* **5,** 337
22. Lundberg, E., and Uhlen, M. (2010) Creation of an antibody-based subcellular protein atlas. *Proteomics* **10,** 3984–3996
23. Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S., Wernerus, H., Björling, L., and Ponten, F. (2010) Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* **28,** 1248–1250
24. Lundberg, E., Fagerberg, L., Klevebring, D., Matic, I., Geiger, T., Cox, J., Algenäs, C., Lundeberg, J., Mann, M., and Uhlen, M. (2010) Defining the transcriptome and proteome in three functionally different human cell lines. *Mol. Syst. Biol.* **6,** 450
25. Nygren, P. A., Eliasson, M., Abrahmsén, L., Uhlén, M., and Palmcrantz, E. (1988) Analysis and use of the serum albumin binding domains of streptococcal protein G. *J Mol Recognit* **1,** 69–74
26. Berglund, L., Björling, E., Jonasson, K., Rockberg, J., Fagerberg, L., Al-Khalili Szigyarto, C., Sivertsson, A., and Uhlén, M. (2008) A whole-genome bioinformatics approach to selection of antigens for systematic antibody generation. *Proteomics* **8,** 2832–2839
27. Larsson, M., Graslund, S., Yuan, L., Brundell, E., Uhlén, M., Höög, C., and Ståhl, S. (2000) High-throughput protein expression of cDNA products as a tool in functional genomics. *J. Biotechnol.* **80,** 143–157
28. Agaton, C., Galli, J., Höidén Guthenberg, I., Janzon, L., Hansson, M., Asplund, A., Brundell, E., Lindberg, S., Ruthberg, I., Wester, K., Wurtz, D., Höög, C., Lundeberg, J., Ståhl, S., Pontén, F., and Uhlén, M. (2003) Affinity proteomics for systematic protein profiling of chromosome 21 gene products in human tissues. *Mol. Cell. Proteomics* **2,** 405–414
29. Li, M. Z., and Elledge, S. J. (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods* **4,** 251–256
30. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expression Purification* **41,** 207–234
31. Skerra, A., and Schmidt, T. G. (2000) Use of the Strep-Tag and streptavidin for detection and purification of recombinant proteins. *Methods Enzymol.* **326,** 271–304
32. Wiśniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nat. Methods* **6,** 359–362
33. Wiśniewski, J. R., Zougman, A., and Mann, M. (2009) Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. *J. Proteome Res.* **8,** 5674–5678
34. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75,** 663–670
35. Olsen, J. V., Schwartz, J. C., Griep-Raming, J., Nielsen, M. L., Damoc, E., Denisov, E., Lange, O., Remes, P., Taylor, D., Splendore, M., Wouters, E. R., Senko, M., Makarov, A., Mann, M., and Horning, S. (2009) A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. *Mol. Cell. Proteomics* **8,** 2759–2769
36. Geiger, T., Cox, J., and Mann, M. (2010) Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Mol. Cell. Proteomics* **9,** 2252–2261
37. Lu, A., Waanders, L. F., Almeida, R., Li, G., Allen, M., Cox, J., Olsen, J. V., Bonaldi, T., and Mann, M. (2007) Nanoelectrospray peptide mapping revisited: Composite survey spectra allow high dynamic range protein characterization without LCMS on an orbitrap mass spectrometer. *Int. J. Mass Spectrom.* **268,** 158–167
38. 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
39. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10,** 1794–1805
40. Matic, I., Jaffray, E. G., Oxenham, S. K., Groves, M. J., Barratt, C. L., Tauro, S., Stanley-Wall, N. R., and Hay, R. T. (2011) Absolute SILAC-Compatible Expression Strain Allows Sumo-2 Copy Number Determination in Clinical Samples. *Journal of proteome research* **10,** 4869–4875