

# Quantitative Proteomics Targeting Classes of Motif-containing Peptides Using Immunoaffinity-based Mass Spectrometry\*<sup>§</sup>

Niclas Olsson‡, Peter James‡, Carl A. K. Borrebaeck‡, and Christer Wingren‡§

The development of high-performance technology platforms for generating detailed protein expression profiles, or protein atlases, is essential. Recently, we presented a novel platform that we termed global proteome survey, where we combined the best features of affinity proteomics and mass spectrometry, to probe any proteome in a species independent manner while still using a limited set of antibodies. We used so called context-independent-motif-specific antibodies, directed against short amino acid motifs. This enabled enrichment of motif-containing peptides from a digested proteome, which then were detected and identified by mass spectrometry. In this study, we have demonstrated the quantitative capability, reproducibility, sensitivity, and coverage of the global proteome survey technology by targeting stable isotope labeling with amino acids in cell culture-labeled yeast cultures cultivated in glucose or ethanol. The data showed that a wide range of motif-containing peptides (proteins) could be detected, identified, and quantified in a highly reproducible manner. On average, each of six different motif-specific antibodies was found to target about 75 different motif-containing proteins. Furthermore, peptides originating from proteins spanning in abundance from over a million down to less than 50 copies per cell, could be targeted. It is worth noting that a significant set of peptides previously not reported in the PeptideAtlas database was among the profiled targets. The quantitative data corroborated well with the corresponding data generated after conventional strong cation exchange fractionation of the same samples. Finally, several differentially expressed proteins, with both known and unknown functions, many relevant for the central carbon metabolism, could be detected in the glucose- versus ethanol-cultivated yeast. Taken together, the study demonstrated the potential of our immunoaffinity-based mass spectrometry platform for reproducible quantitative proteomics targeting classes of motif-containing peptides. *Molecular & Cellular Proteomics* 11: 10.1074/mcp.M111.016238, 342–354, 2012.

In the quest for disease-associated biomarkers, the deciphering of the human proteome(s) will be central (1). Albeit powerful mass spectrometry (MS)-based technology platforms have been developed and frequently applied (2–4), the output in terms of validated biomarkers have so far been limited, mainly due to technological issues (5, 6). In recent years, affinity proteomics based on antibody microarrays have become an established proteomic technology for protein expression profiling of complex proteomes (7–11). To date, the technology has been applied in several clinical applications, demonstrating its potential for, e.g. biomarker discovery, improved diagnosis and prognosis, as well as classification (8, 12–15). Despite the success, the possibility to run large-scale and, in particular, discovery mode projects using the conventional antibody array designs have been limited (7, 8, 16). First, antibodies of only known specificities, *i.e.* directed against preselected targets of a known specie, have been included, thereby excluding the possibility to discover novel targets (across species). Second, the resolution of an antibody array is directly related to the sheer number of antibodies included and their range of specificities, which tends to be a bottleneck. Third, the numbers of readily available antibodies, designed for microarray applications, have per se been limited.

To bypass these technological hurdles and advance further, and even to provide quantitative capabilities, the most attractive features of affinity proteomics and MS could be combined (17, 18). The biological sample would then be digested and exposed to peptide-specific antibodies, after which any specifically enriched peptides would be detected, identified, and potentially quantified using MS. This was demonstrated in the stable isotope standard capture with antipeptide antibodies set-up, but, like conventional antibody arrays, this platform relied on the use of one binder per unique peptide/protein (19, 20). To circumvent the need of having to generate numerous antibodies, we (18) and others (17), have recently presented the novel concept of using antibodies directed against short peptide-motifs (epitopes) shared among up to hundreds of different peptides/proteins. This would provide an inherent capability of probing any proteome in a discovery mode, in a species independent manner, while still using a very limited number of antibodies. Based on this concept, we have recently designed a platform denoted global proteome survey

From the ‡Department of Immunotechnology, Lund University, Lund, Sweden, CREATE Health, BMC D13, Lund, Sweden

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

Received November 28, 2011, and in revised form, March 30, 2012

Published, MCP Papers in Press, April 27, 2012, DOI 10.1074/mcp.M111.016238

(GPS)<sup>1</sup>, based on human recombinant single-chain fragment variable (scFv) antibodies (18, 22), while Joos *et al.* presented the triple X proteomic (TXP) set-up, relying on polyclonal and/or monoclonal antibodies. In our case, the scFv antibodies, microarray adapted by molecular design (7, 8, 23), were derived from a large phage-display library (24), representing a renewable probe source displaying an extensive range of specificities. In this manner, one hundred of such scFv antibodies, denoted context-independent motif specific (CIMS) antibodies, could theoretically cover almost 50% of the non-redundant human proteome (18, 22), a concept supported by a recent *in-silico* motif study of the human proteome (25). Recently, we demonstrated experimental proof-of-principle for that a limited number of CIMS antibodies could be used to profile crude, digested proteomes when combined with a mass spectrometry-based read-out (22). The GPS platform was in fact shown to provide novel and broad coverage, and to have the potential to reach deep into a proteome in a species independent manner.

In this study, we have demonstrated the GPS set-up with respect to its quantitative capability, reproducibility, sensitivity, and coverage, by using the stable isotope labeling by amino acids in cell culture (SILAC) approach (26), targeting SILAC-labeled yeast cultivated in either glucose or ethanol. In order to evaluate the quantitative ability in more detail, the same samples were in parallel experiments fractionated using conventional strong cation exchange chromatography (SCX). The data showed the potential and applicability of our immunoaffinity-based mass spectrometry GPS platform for reproducible quantitative proteomics targeting classes of motif-containing peptides.

#### EXPERIMENTAL PROCEDURES

**Production of CIMS Antibodies**—Six human recombinant CIMS scFv antibodies (clones 1-B03, 15-A06, 17-E02, 32-3A-G03, 33-3D-F06, and 34-3A-D10) directed against six short C-terminal amino acid peptide motifs (denoted M-1, M-15, M17, M-32, M-33, and M-34), were selected from the n-CoDeR library (24), and kindly provided by Biolnvent International AB, Lund, Sweden (supplemental Table S1). The specificity and dissociation constant (sub  $\mu\text{M}$  range) of the CIMS antibodies have recently been evaluated elsewhere (22) (Olsson *et al.*, ms in prep.). Noteworthy, the antibody selection procedure adopted so far, was not designed to generate binders with a strict 4 or 6 amino acid motif specificity, why the sequence of the bound motif varied (wobbled). In general, only two or three positions appeared to be more critical (“anchor residues”), whereas a higher degree of wobbling was allowed in the other positions. The antibodies were produced in 100 ml *E. coli* cultures and purified using affinity

chromatography on Ni<sup>2+</sup>-NTA agarose (Qiagen, Hilden, Germany). Bound molecules were eluted with 250 mM imidazole, dialyzed against phosphate-buffered saline (PBS) (pH 7.4) for 72 h and then stored at +4 °C until further use. The protein concentration was determined by measuring the absorbance at 280 nm using a Nano-drop-1000. The integrity and purity of the scFv antibodies was evaluated by running Protein 80 chips on Agilent Bioanalyzer (Agilent, Waldbronn, Germany).

**Coupling of CIMS-scFv Antibodies to Magnetic Beads**—Purified scFvs were individually coupled to magnetic beads (M-270 carboxylic acid-activated, Invitrogen Dynal, Oslo) using manufacture provided protocols with some minor modifications as described before (22). Briefly, 180–250  $\mu\text{g}$  purified scFv was covalently coupled (EDC-NHS chemistry) to  $\sim 9$  mg (300  $\mu\text{l}$ ) of magnetic beads, and stored in 0.005% (v/v) Tween-20 in PBS at 4 °C until use.

**Yeast Culture**—*Saccharomyces cerevisiae* was chosen as model organism in order to represent a well-characterized complex proteome. To this end, the yeast strain YAL6B, containing deficient genes for both lysine and arginine synthesis, and kindly provided by Prof. O.N. Jensen (University of Southern Denmark, Odense, Denmark), was applied (27). A preculture was started in yeast nitrogen base medium (YNB) containing 0.055 g/L adenine, 0.055 g/L tyrosine, 0.045 g/L uracil and supplemented with 2% (w/v) glucose and necessary amino acids. Thirty mg/L arginine and 30 mg/L lysine were added to the medium. The culture was incubated overnight at 30 °C on a rotary shaker at 200 rpm, followed by dilution to an OD<sub>600</sub> of 1. Then, 1 ml of the overnight culture was transferred to either normal or heavy YNB medium supplemented with 30 mg/L <sup>13</sup>C<sub>6</sub>, <sup>15</sup>N<sub>4</sub>-arginine and 30 mg/L <sup>13</sup>C<sub>6</sub>-lysine (Cambridge Isotope Laboratories, Andover, MA). The cultures were incubated at 30 °C and grown overnight to an OD<sub>600</sub> of 1 in order to ensure complete incorporation of isotopic amino acids. Finally, two yeast cell cultures were then initiated with either 2% (w/v) glucose or 0.05% (w/v) glucose and 3% (v/v) ethanol as the carbon source in YNB medium supplemented with necessary amino acids. Swap experiments were performed as a biological replicate in order to have the isotopic labeled amino acids present in both growth conditions. Hence, in total four cultures were initiated (glucose, glucose (K<sup>6</sup>, R<sup>10</sup>), ethanol, and ethanol (K<sup>6</sup>, R<sup>10</sup>)). In addition, two separate cultures (glucose and glucose (K<sup>6</sup>, R<sup>10</sup>)) were initiated for the spike-in proteome ratio experiments. The cultures were grown in 30 °C on a rotary shaker at 200 rpm and harvested in log phase at OD<sub>600</sub> 0.7. The cultures were centrifuged at 5000 rpm (10 min at 4 °C). The pellets were dissolved in 10 ml ice-cold Milli-Q water and then centrifuged at 2000  $\times$  g (5 min at 4 °C). In order to preserve the samples from degradation (28), the pellets were resuspended in 1.5 ml cold 10% (v/v) TCA, transferred to 2 ml tubes and finally collected by centrifugation at 13,200 rpm for 1 min at 4 °C. Samples were snap-frozen in liquid nitrogen and stored at –80 °C until further use.

**Protein Extraction and Sample Preparation**—Cell pellets were thawed on ice, resuspended in 500  $\mu\text{l}$  extraction buffer 8 M Urea, 30 mM Tris, 5 mM MgAc and 4% (w/v) CHAPS (pH 8.5). Equal amount (500  $\mu\text{l}$ ) of resuspended cells from the different growth conditions of ethanol and glucose were pooled (heavy + light) and 1.0 g glass beads (0.55 mm) added to the pooled cells. For the separately planned spike-in proteome ratio experiments, no pooling of resuspended cells was done, and protein extracts were isolated separately allowing for downstream mixing of various “heavy” (H) and “light” (L) proteome amounts. Extractions were performed using a BeadBeater with 8  $\times$  60 s mixing incubations at 4 °C (and 60s intervals on ice between mixing steps) followed by centrifugation at 10,000  $\times$  g (10 min at 4 °C). The buffer was exchanged to 0.15 M HEPES, 0.5 M Urea using Zeba desalting spin columns (Pierce, Rockford, IL), before the protein concentration was determined using Total Protein Kit, Micro Lowry (Sigma, Saint Louis, MO). The average protein concentrations

<sup>1</sup> The abbreviations used are: GPS, global proteome survey; CIMS, context independent motif specific; EDC, N-Ethyl-N’-(3-dimethylaminopropyl)carbodiimide; H/L, ratio heavy to light; LTQ, linear ion trap; MRM, multiple reaction monitoring; scFv, single-chain Fragment variable; SCX, strong cation exchange chromatography; SILAC, stable isotope labeling by amino acids in cell culture; TCA, tricarboxylic cycle; TCEP-HCl, Tris(2-Carboxyethyl) phosphine Hydrochloride; TXP, triple X proteomics.

(triplicates) were determined to be 2.03  $\mu\text{g}/\mu\text{l}$  for the yeast grown in glucose and 1.98  $\mu\text{g}/\mu\text{l}$  for the yeast grown in glucose in the presence of isotopic amino acids. Finally, the samples were aliquoted and stored at  $-80^\circ\text{C}$  until further use.

For the quantitative proteome spike-in control GPS-experiments, different amounts of protein extracts from glucose grown yeast (unlabeled and SILAC labeled) were mixed at three H and L ratios (20/80 H/L, 50/50 H/L, and 80/20 H/L) based on the total protein concentrations. The protein extracts were thawed, reduced, alkylated, and trypsin digested. First, 0.05% (w/v) SDS and 5 mM TCEP-HCl (Thermo Scientific, Rockford, IL) were added, and the samples were reduced for 60 min at  $56^\circ\text{C}$ . The samples were cooled to room temperature before iodoacetamide was added to 10 mM and then alkylated for 30 min at room temperature. Next, sequencing-grade modified trypsin (Promega, Madison, WI) was added at 20  $\mu\text{g}$  per mg of protein for 16 h at  $37^\circ\text{C}$ . In order to ensure complete digestion, a second aliquot of trypsin (10  $\mu\text{g}$  per mg protein) was added and the tubes were incubated for an additional 3 h at  $37^\circ\text{C}$ . Finally, the digested samples were aliquoted and stored at  $-80^\circ\text{C}$  until further use. The addition of a trypsin inhibitor directly after thawing the digested sample, did not apparently affect the subsequent GPS analysis (Olsson *et al.*, unpublished observations).

**GPS Capture of Tryptic Digests**—A standardized sample work-flow was adopted. For each capture experiment, a 35  $\mu\text{l}$  solution of pre-washed (300  $\mu\text{l}$  PBS) CIMS-conjugated beads was used. A tryptic digest (20  $\mu\text{g}$  was thawed just prior to use and diluted in PBS) into a final incubation volume of 30–35  $\mu\text{l}$  and then incubated with the beads for precisely 15 min with gentle mixing. Next, the tubes were placed on a magnet, the supernatant removed, and the beads were washed twice with 62 and 50  $\mu\text{l}$  PBS, respectively (a total of 5 min) (the beads were transferred to new tubes in between each wash step). Finally, the beads were incubated with 7.5  $\mu\text{l}$  5% (v/v) acetic acid for 1 min in order to elute captured peptides. Two technical replicates were performed for each sample and binder. The eluate was then used directly for mass spectrometry analysis without any additional cleanup.

**SCX Peptide Fractionation**—For SCX fractionation of peptides, 100  $\mu\text{g}$  of each biological sample (ethanol-glucose mix) was concentrated to a volume of 100  $\mu\text{l}$  by speed-vac, followed by an addition of 500  $\mu\text{l}$  5 mM  $\text{KH}_2\text{PO}_4$ , 25% (v/v) acetonitrile (ACN) (pH 2.9). The SCX fractionation was performed on an ICAT-Cartridge (Applied Biosystems, Foster City, CA) cation exchange. The resin was first activated by adding 5 mM  $\text{KH}_2\text{PO}_4$ , 25% (v/v) ACN, 1 M KCl (pH 2.9) and then rinsed with 5 mM  $\text{KH}_2\text{PO}_4$ , 25% (v/v) ACN (pH 2.9) before the sample was loaded. An eight-step elution gradient from 30 to 500 mM KCl was used. Each 500  $\mu\text{l}$  fraction (in total 8), was dried down and re-suspended in 5% (v/v) formic acid followed by a c18 clean up performed with UltraMicroSpinColumns (Vydac C18 Silica from The Nest Group, Southborough, MA). The wetting of the C18 Silica was performed by adding  $2 \times 50 \mu\text{l}$  70% ACN, 5% (v/v) formic acid followed by equilibration by adding  $2 \times 50 \mu\text{l}$  5% (v/v) formic acid. The sample was loaded and washed with  $4 \times 50 \mu\text{l}$  5% (v/v) formic acid. Finally, the samples were eluted with 50  $\mu\text{l}$  50% (v/v) ACN, 5% (v/v) formic acid, dried down and resuspended in 30  $\mu\text{l}$  0.1% (v/v) formic acid.

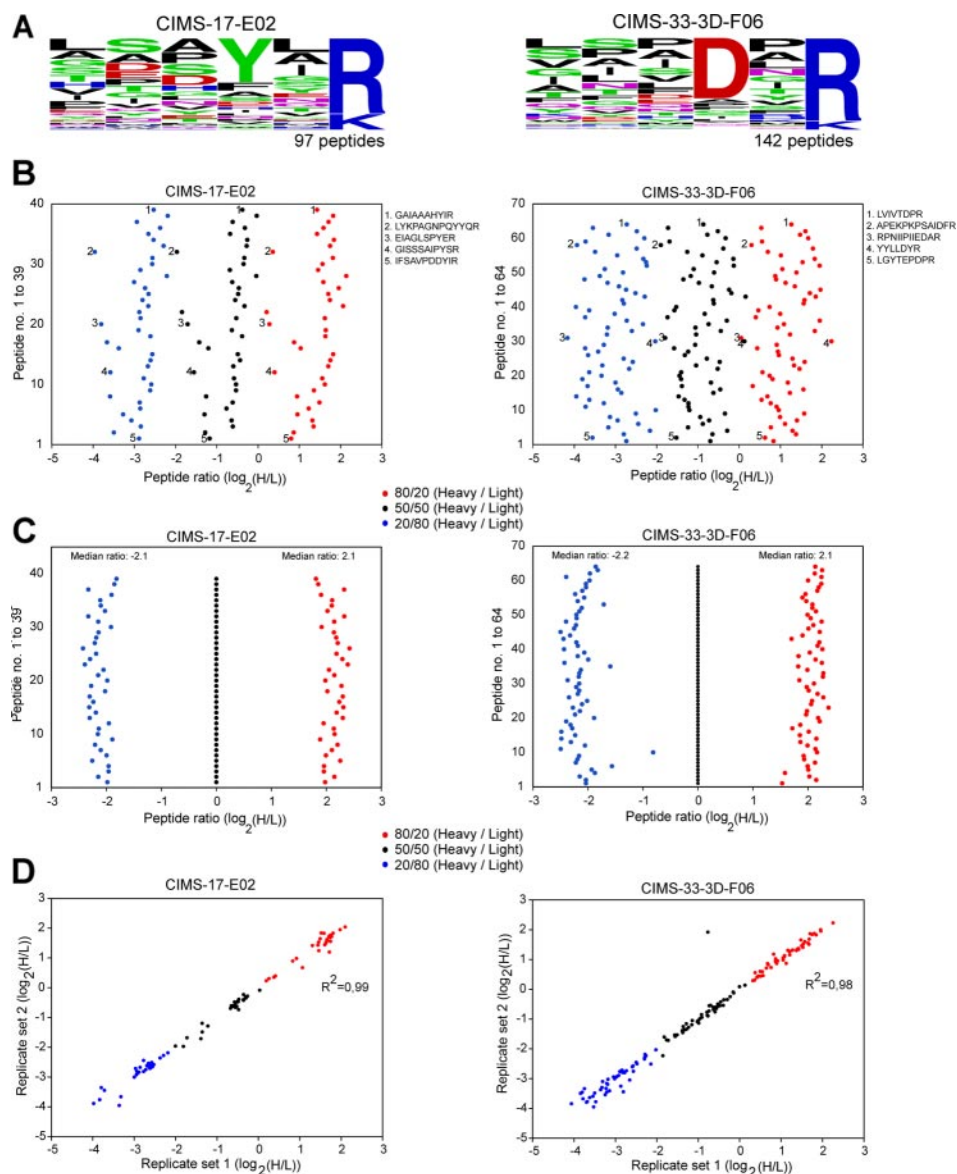
**Mass Spectrometry Analysis**—An ESI-LTQ-Orbitrap XL (Thermo Electron, Bremen, Germany) coupled to an Eksigent 2D nano HPLC (Eksigent technologies, Dublin, CA) was used for all samples analyzed. The auto-sampler injected 6  $\mu\text{l}$  of GPS-generated eluates and 5  $\mu\text{l}$  of the SCX fractions respectively. The peptides were trapped on a precolumn (Zorbax 300SB-C18  $5 \times 0.3 \text{ mm}$ , 5  $\mu\text{m}$ , Agilent) and separated on a reversed-phase analytical column (Zorbax 300SB-C18  $150 \times 0.75 \text{ mm}$ , 3.5  $\mu\text{m}$ , Agilent). The flow rate was 350 nl/min. Solvent A consisted of 0.1% (v/v) formic acid in water and solvent B of 0.1% (v/v) formic acid in acetonitrile. The total runtime was 70 min

starting with a 5 min wash of the GPS generated peptides. For the SCX fractions, the total run time was 80 min starting with a 15 min wash of the peptides. The linear ion trap (LTQ)-Orbitrap was operated in data-dependent mode to automatically switch between Orbitrap-MS and LTQ-MS/MS acquisition. Survey full scan MS spectra (from  $m/z$  400 to 2000) were acquired in the Orbitrap with a resolution of 60,000 at  $m/z$  400 using the lock mass option for internal calibration. The seven most intense ions with charge state 2 and up were sequentially isolated for CID-fragmentation in the LTQ with a normalized collision energy of 35% and the resulting fragment ions were recorded in the LTQ. For the proteome spike-in control experiments (20/80 H/L, 50/50 H/L, and 80/20 H/L), using CIMS-17-E02 and CIMS-33-3D-F06, a total of 12 LC-MS/MS runs were performed (2 capture experiments per binder and proteome mix). The mixed ratio (e.g. 20/80) was based on the total protein concentration (see above). In order to evaluate any nonspecific background binding peptides, blank beads, *i.e.* beads without any conjugated antibody, were exposed to a 50/50 H/L proteome mixture (supplemental Table S2A-C). Notably, only 13 background peptides, corresponding to 13 proteins, were identified when the eluates from two blank runs were analyzed. For the ethanol-glucose mixed proteome profiling experiments a total of 24 LC-MS/MS runs were performed with the GPS setup (two capture experiments per binder on two biological samples). In total, 32 LC-MS/MS runs were performed on the ethanol-glucose mixed proteomes of the SCX fractionated samples (two runs (technical replicates) for each of the eight eluted fractions from the two biological samples). The total number of identified peptides and proteins were lower in biosample 2 (isotopic amino acids present in glucose condition) than biosample 1, but as this was noticed for both the GPS and SCX set-up, this was attributed to the sample and not the experimental set-ups.

**Data Processing and Analysis**—Raw Orbitrap full scan MS and MS/MS spectra were processed by MaxQuant (v 1.0.13.13) as described (29) using default settings and MASCOT Daemon (v 2.2.2 Matrix Science, London UK) as the database search engine for peptide identifications. Briefly, MS/MS peak lists were filtered to contain at most six peaks per 100 Da interval and searched by Mascot against a forward and a random combined database (*Saccharomyces cerevisiae* Swiss-Prot, 13-Oct 2009 with 6890 forward sequences, and with a random generated decoy version with same amino acid distribution and length, resulting in a total of 13,780 sequences). Tandem mass spectra were matched with a mass tolerance of 7 ppm on precursor mass and 0.5 Da for fragment ions. Carbamidomethylation of cysteine was selected as a fixed modification, and oxidation of methionine and acetylation of the protein N terminus were used as variable modifications. Labeled arginine and lysine were specified as fixed or variable. A false discovery rate of 0.01 was used (estimated on the basis of the number of identified random hits). The tool WebLogo was used to generate the motif figures for selected binders enriched sequence motifs (30). The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) tool version 8.3 (31) was used for creating protein-protein interaction (PPI) evidence maps with default settings (medium significance) for a subset of proteins identified as differentially expressed.

**Data**—The data associated with this study is downloadable from Proteome Commons (<http://proteomecommons.org/>) Tranche using the following hash: xNBoKsLK7paM+4ZEI3nSPo7TTVaZw4TrVyZt v22cgMiX8fW+km9Bky96GrXoduVjmWCBO1mmVSeEFKWvUZLU NlvuAzYAAAAAABb6g== and the following passphrase: pn-CaYnzyvfitGXtjnQqy. Included are all raw files that were processed by MaxQuant as described above, identified proteins, sequences of identified peptides and peptide evidences (*i.e.* SILAC ratio measurements), annotated MS/MS spectra for single peptide identifications, and the yeast database file used for Mascot searches.





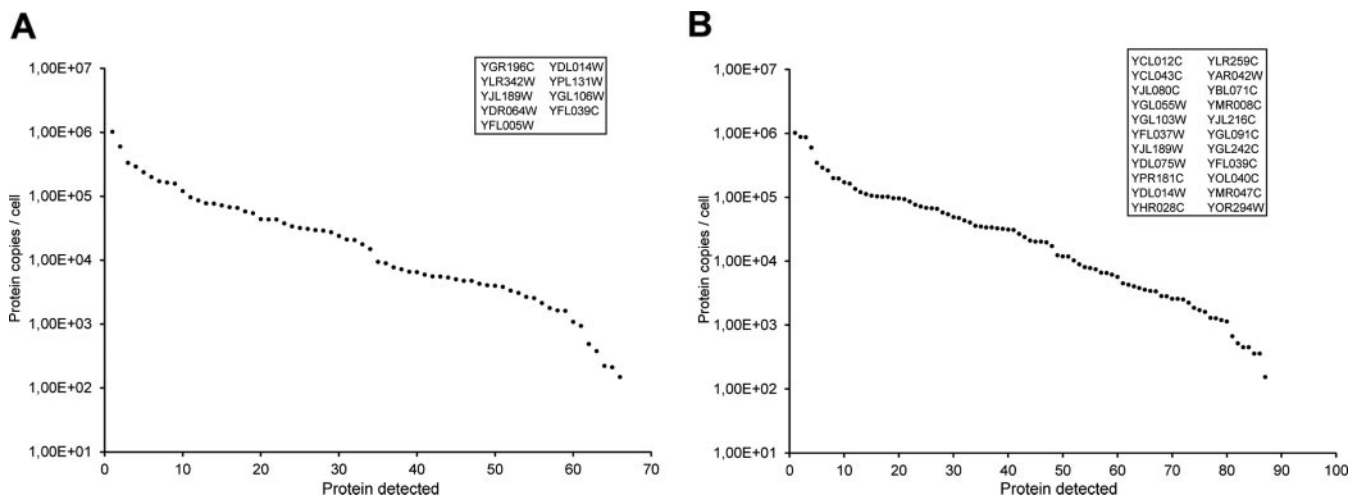
**FIG. 1. Quantitative accuracy and reproducibility.** *A*, Experimentally refined binding motif amino acid sequences as compiled from all MS-MS detected captured peptides (excluding background peptides) in the spike-in proteome experiments (glucose grown SILAC-labeled yeast proteomes mixed at three ratios of H and L) for two CIMS antibodies, clones CIMS17-E02 (97 peptides) and CIMS-33-3D-F06 (142 peptides). The frequency of each individual residue in the last six C-terminal positions is indicated. *B*, Peptide intensity/ratio distribution for peptides enriched, from predefined proteome mixtures, by the two CIMS antibodies. Data limited to YXR motif-containing peptides for CIMS-17-E02 and DXR motif-containing peptides for CIMS-33-3D-F06. (The two replicate captures for each condition were analyzed as one experiment in MaxQuant. Hence, raw (non-normalized) ratios, as calculated from associated redundant peptide ratios by means of median of the two replicates, are displayed. Furthermore, data limited to peptides with MaxQuant reported ratios in all conditions and replicate capture trials resulting in 39 different peptides for CIMS-17-E02 and 64 different peptides for CIMS-33-3D-F06). Noteworthy, the peptides were ordered based on total intensity, resulting in that peptide number 39 and 64 were the two peptides observed with the highest total intensity. *C*, The same data as in panel *B*, but now compensated (normalized) against the corresponding 50/50 ratio measurement set to be 0. *D*, Reproducibility between replicate capture experiments for CIMS-17-E02 (limited to YXR motif-containing peptides) and CIMS-33-3D-F06 (limited to DXR motif-containing peptides).

## RESULTS

In this study, we have demonstrated the efficacy of the GPS technology platform for quantitative proteomics. To this end, SILAC-labeled protein extracts from *S. cerevisiae*, cultivated in glucose and ethanol, were used as model proteomes. An overall

workflow outlining all steps of the set-up is schematically shown in supplemental Fig. S1.

**Quantitative Accuracy**—To investigate the quantitative accuracy and reproducibility of the GPS measurements, two CIMS antibodies, clones CIMS-17-E02 and CIMS-33-3D-



**Fig. 2. Dynamic range and sensitivity.** Glucose grown yeast SILAC-proteomes, mixed at three ratios of H and L, were profiled using two CIMS antibodies, A, CIMS-17-E02 and B, CIMS-33-3D-F06, and mapped to the known protein abundances generated by orthogonal methods (32). In total, the data was based on six capture trials (two for each isotopically labeled spike in proteome ratio mix) per CIMS-antibody (supplemental Tables S3–S5). For Swiss-Prot ID's matching two different synonymous yeast SGD ids and two different values in (32), the highest reported value was plotted. In cases with multiple Swiss-Prot ID's, they were checked in alphabetical order for presence in (32), and if no value was present the second Swiss-Prot id was checked and onwards. A number of identified peptides/proteins were reported to be present in less than 50 copies/cell or had no reported value at all and thereby could not be plotted and are listed separately in a box.

F06, were used to profile glucose grown yeast SILAC-proteomes mixed at three known ratios of H and L (80/20, 50/50, and 20/80) (Fig. 1 and supplemental Tables S3 and S4). In total, two captures per binder and proteome mix, *i.e.* 12 runs, were performed. First, the sequence of the enriched peptides were determined and matched to the motif specificity of the CIMS antibodies. The results showed that a total of 100 different peptides, including three background peptides, were enriched by CIMS-17-E02, whereas 145 different peptides (three background peptides) were captured by CIMS-33-3D-F06 (Fig. 1A).

Next, the H/L ratios were determined for all the enriched peptides (Fig. 1B and supplemental Figs. S2A and S2B). In Fig. 1B, representative data, illustrated for YXR (CIMS-17-E02) and DXR (CIMS-33-3D-F06) motif-containing peptides are plotted using non-normalized raw data. The data showed that the peptides were grouped into three distinct clusters, with the expected differences in peptide ratios between the three groups. Hence, the data outlined the quantitative accuracy of the GPS method. The fact that the observed ratios, 1.6 (20/80),  $-0.6$  (50/50), and  $-2.8$  (80/20) for CIMS-17-E02, and 1.1,  $-0.9$ , and  $-2.9$  for CIMS-33-3D-F06 (Fig. 1B), were relatively displaced as compared with the expected ratios ( $-2$ , 0, and 2) could be explained by that non-normalized data was used. When the data was normalized toward the 50/50 ratio, the 20/80 and 80/20 ratios were clearly distributed around the anticipated values, displaying median ratios around  $-2.1$  and  $2.1$ , respectively (Fig. 1C and Supplemental Fig. S2B).

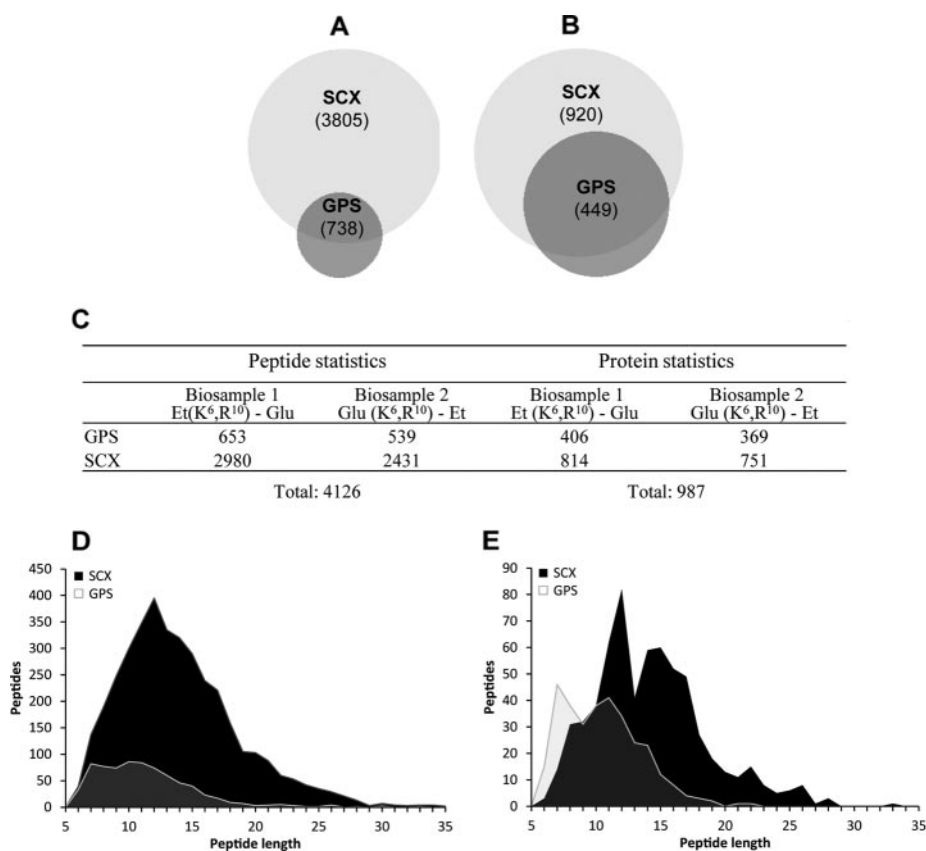
The reproducibility of the technical replicates, *i.e.* encompassing the entire GPS set-up including the capture step, was then determined (supplemental Fig. S2C). Representative data, illustrated for YXR (CIMS-17-E02) and DXR (CIMS-33-

3D-F06) motif-containing peptides, is plotted in Fig. 1D. The results showed that high  $r^2$ -values (0.98 and 0.99) were obtained, and the total median percent difference, considering all six captures and analytes, was 5.9% to 6.8% (supplemental Fig. S2C).

The technical reproducibility was also determined for analytes displaying a significantly larger fold change than the artificial mixtures used above ( $\leq$ fourfold changes), by targeting SILAC-proteomes generated from *S. cerevisiae* cultivated in glucose or ethanol. Representative data, illustrated for CIMS antibody CIMS-33-3D-F06, is shown in supplemental Fig. S3. After removing a single potential outlier peptide, present in both biological replicates,  $r^2$ -values between 0.96 and 0.98 were observed.

**Dynamic Range**—In order to evaluate the dynamic range (and sensitivity) of the GPS platform, glucose grown yeast SILAC-proteome were profiled using two CIMS antibodies, and the identified yeast proteins were then mapped to the known absolute protein abundances generated by orthogonal methods (32) (Fig. 2 and supplemental Tables S5A and S5B). For both CIMS antibodies, the results showed that yeast proteins could be detected in a wide range of abundances, spanning over a million and reaching down to less than 50 copies per cell (Fig. 2). Noteworthy, a significant set of the identified peptides (19 and 27%, respectively) had not previously been reported in PeptideAtlas (supplemental Tables S3B and S4B). Hence, the GPS set-up was shown to display a dynamic range of at least three orders of magnitude, high sensitivity, and a complementary peptide coverage.

**Coverage—Comparison of GPS and SCX**—To further evaluate the quantitative capability of GPS, the data was com-



**FIG. 3. Coverage—comparison of GPS and SCX.** Yeast SILAC-proteomes, cultivated in ethanol or glucose, were profiled using GPS and SCX. *A*, Venn diagram showing the number of peptides identified by GPS (based on 6 CIMS antibodies) and SCX; a total number of 4126 peptides were identified of which 417 were detected in both. *B*, Venn diagram showing the number of proteins identified by GPS and SCX; a total number of 987 proteins were identified of which 382 were detected in both. *C*, Statistics of all the identified peptides and proteins. *D*, peptide length plotted for all identified peptides using GPS and SCX. *E*, peptide length plotted for peptides identified in both biological replicates with a ratio count of  $\geq 2$  using GPS and SCX.

pared with that obtained for the same samples analyzed in parallel experiments using one of the main orthogonal approaches within conventional quantitative proteomics, namely SCX LC-MS/MS (Fig. 3 and supplemental Tables S6A–S6C). To this end, SILAC-proteomes generated from *S. cerevisiae*, cultivated in glucose or ethanol, were profiled using six CIMS antibodies and any differentially expressed analytes were identified and quantified. Although two biological replicates and two technical replicates were applied in both cases, the GPS set-up relied on six CIMS antibodies (in total 24 runs) and the SCX set-up on eight fractions (in total 32 runs).

The GPS platform was found to identify 738 peptides (Fig. 3A and supplemental Table S6B) corresponding to 449 proteins (Fig. 3B and supplemental Table S6A), *i.e.* in average detecting 1.6 peptides/protein. Each CIMS antibody was in average found to target 75 motif-containing but different proteins. Notably, the cross-talk was found to be low (supplemental Fig. S4). In more detail, 80% of all peptides were only captured by one of the CIMS antibodies, whereas 12% by two CIMS antibodies, and single % by three or more CIMS bind-

ers. In comparison, the SCX set-up was shown to detect 3805 peptides (Fig. 3A and supplemental Table S6A) corresponding to 920 proteins (Fig. 3B and supplemental Table S6A), in average detecting 3.8 peptides/protein. Notably, 321 peptides and 67 proteins were uniquely detected with the GPS methodology, and 3388 peptides and 539 proteins with SCX, respectively. Hence, the GPS and SCX set-ups were found to display overlapping and complementary protein expression profiling data. This notion was further supported by the observation that GPS and SCX were found to target peptides of overlapping but different average lengths. More specifically, a median peptide length of 11 (GPS) *versus* 13 (SCX) amino acids was observed based on all peptides (Fig. 3D), and 10 *versus* 14 amino acids (Fig. 3E) when limited to peptides identified and quantified in both biological replicates (ratio count  $\geq 2$ ).

**Quantitative Data—Comparison of GPS and SCX**—Next, we determined the correlation between the quantitative data generated using GPS and SCX, again targeting SILAC-proteomes generated from *S. cerevisiae*, cultivated in glucose or ethanol (Fig. 4). The results showed that a high correlation

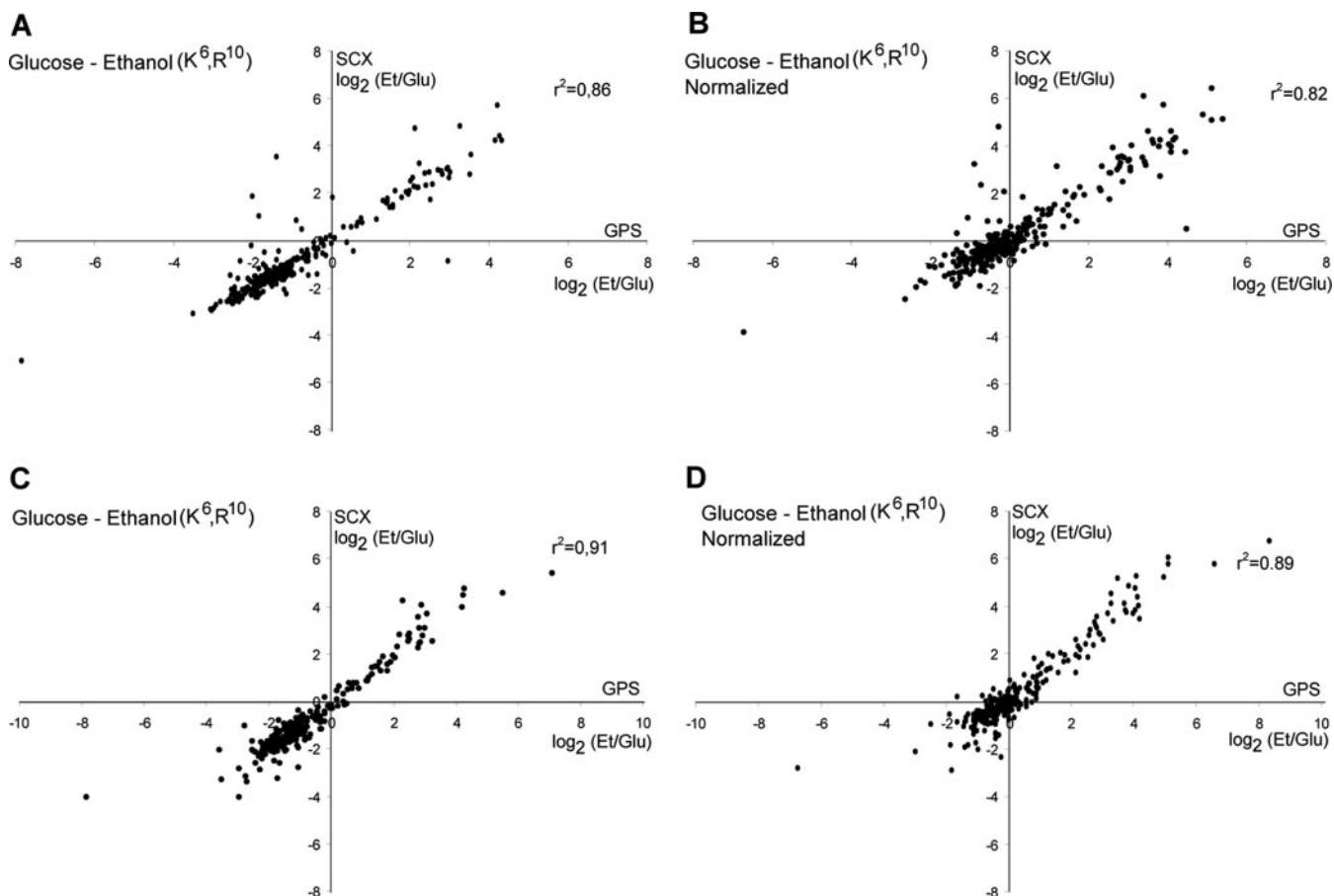


FIG. 4. **Quantitative data—comparison of GPS and SCX.** SILAC-labeled yeast proteomes, cultivated in ethanol or glucose, were profiled using GPS (based on six CIMS antibodies) and SCX. The correlation between the data sets was examined. *A*, all identified peptides for biosample 1 (isotopic amino acids present in ethanol condition). *B*, normalized data for all identified peptides in biosample 1 (isotopic amino acids present in ethanol condition). *C*, all identified proteins for biosample 1 (isotopic amino acids present in ethanol condition). *D*, normalized data for all identified proteins in biosample 1 (isotopic amino acids present in ethanol condition).

between the quantitative data was obtained on both the peptide ( $r^2$ -values of 0.86 and 0.82) and protein ( $r^2$ -values of 0.91 and 0.89) levels covering a large concentration span of various analytes, demonstrating the quantitative capability of GPS.

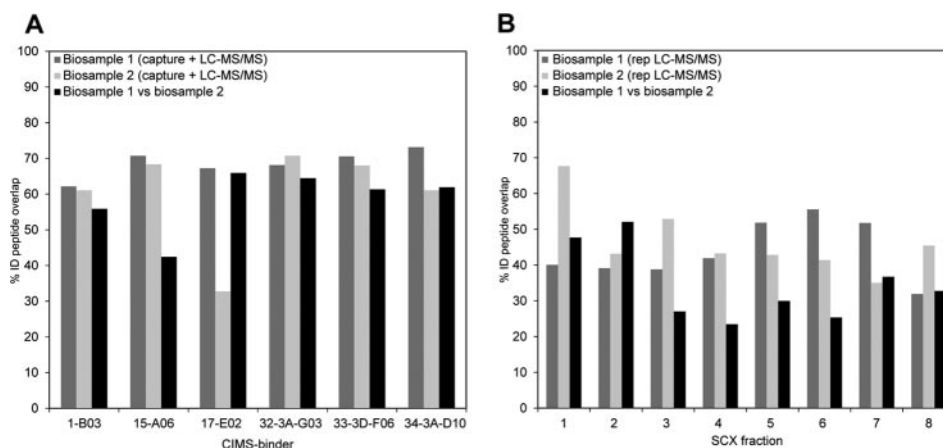
Noteworthy, the reproducibility between the two biological replicates (including swap of isotopic labeled amino acids) for GPS and SCX were then also determined and compared (supplemental Fig. S5). In the case of GPS, the data showed that 319 of 738 (43%) detected peptides were consistently quantified between the biological replicates (supplemental Fig. S5A), although only 641 of 3805 (17%) detected peptides were consistently quantified with SCX (supplemental Fig. S5C). After removing any potential outliers ( $\leq 3$ ), an  $r^2$  correlation of 0.76 was observed for GPS (supplemental Fig. S5B), but only 0.60 for SCX (supplemental Fig. S5D), clearly demonstrating that GPS generated reproducible and stable peptide quantification for biological replicates (including swapping of isotopic amino acids).

**Identification Repeatability—Comparison of GPS and SCX—**Next, the reproducibility of the MS/MS identification was eval-

uated for both GPS and SCX (Fig. 5). To this end, the identification overlap was determined for all replicate runs. The results showed that the median identification overlap was 68% for GPS (Fig. 5A), but only 43% for SCX (Fig. 5B). Furthermore, the identification overlap between biological samples was also found to be significantly higher for GPS than for SCX, 62% versus 31%, respectively (Fig. 5). It should be noted that the GPS data was based on two separate captures and LC-MS/MS runs, *i.e.* technical variation for the entire set-up, while the SCX data was based on the same eluate injected twice for LC-MS/MS analyses, *i.e.* technical variation only for the LC-MS/MS runs.

**Biological Relevance of Identified Proteins for Carbon and Ethanol Metabolism—**Finally, the biological relevance of proteins pin-pointed as differentially expressed in glucose versus ethanol cultivated yeast, using GPS and SCX was examined (Fig. 6 and supplemental Fig. S6). The results showed that 27 and 50 differentially expressed proteins ( $p < 0.01$ , biosample 1) were identified using GPS (Fig. 6A) and SCX (Fig. 6B), respectively. Focusing on the central carbon metabolism





**FIG. 5. Identification repeatability—comparison of GPS and SCX.** The reproducibility of the MS/MS identification was evaluated by determining the identification overlap for all replicate runs. *A*, percentage peptide ID overlap in replicate technical (capture + LC-MS/MS) and biological replicate GPS-assays (based on six CIMS antibodies) (biosample 1, ethanol ( $K^6$ ,  $R^{10}$ ) - glucose; biosample 2, ethanol - glucose ( $K^6$ ,  $R^{10}$ )). It should be noted that the generated values were based on separate capture and LC-MS/MS steps. *B*, percentage peptide ID overlap for replicate runs (same SCX-fraction analyzed twice with LC-MS/MS) and biological replicate runs for the eight generated SCX fractions.

pathways, including the glycolysis, tricarboxylic acid (TCA) cycle, and glyoxylate cycle, several key differentially expressed proteins could be mapped (Fig. 6C). Noteworthy, the GPS set-up and the SCX approach showed coherent profiles.

In the case of GPS, the nature of the differentially expressed proteins was further characterized by generating protein-protein interaction maps (Fig. 6D). The results showed that several of the up-regulated proteins were involved in the carbon metabolism pathways and energy production, whereas many of the down-regulated proteins were directly or indirectly involved in transcription and translation, such as ribosomal proteins, ARC1, and TIF3. Taken together, the data demonstrated the applicability of GPS for quantitative proteomics in a reproducible and sensitive manner targeting biologically relevant proteins.

#### DISCUSSION

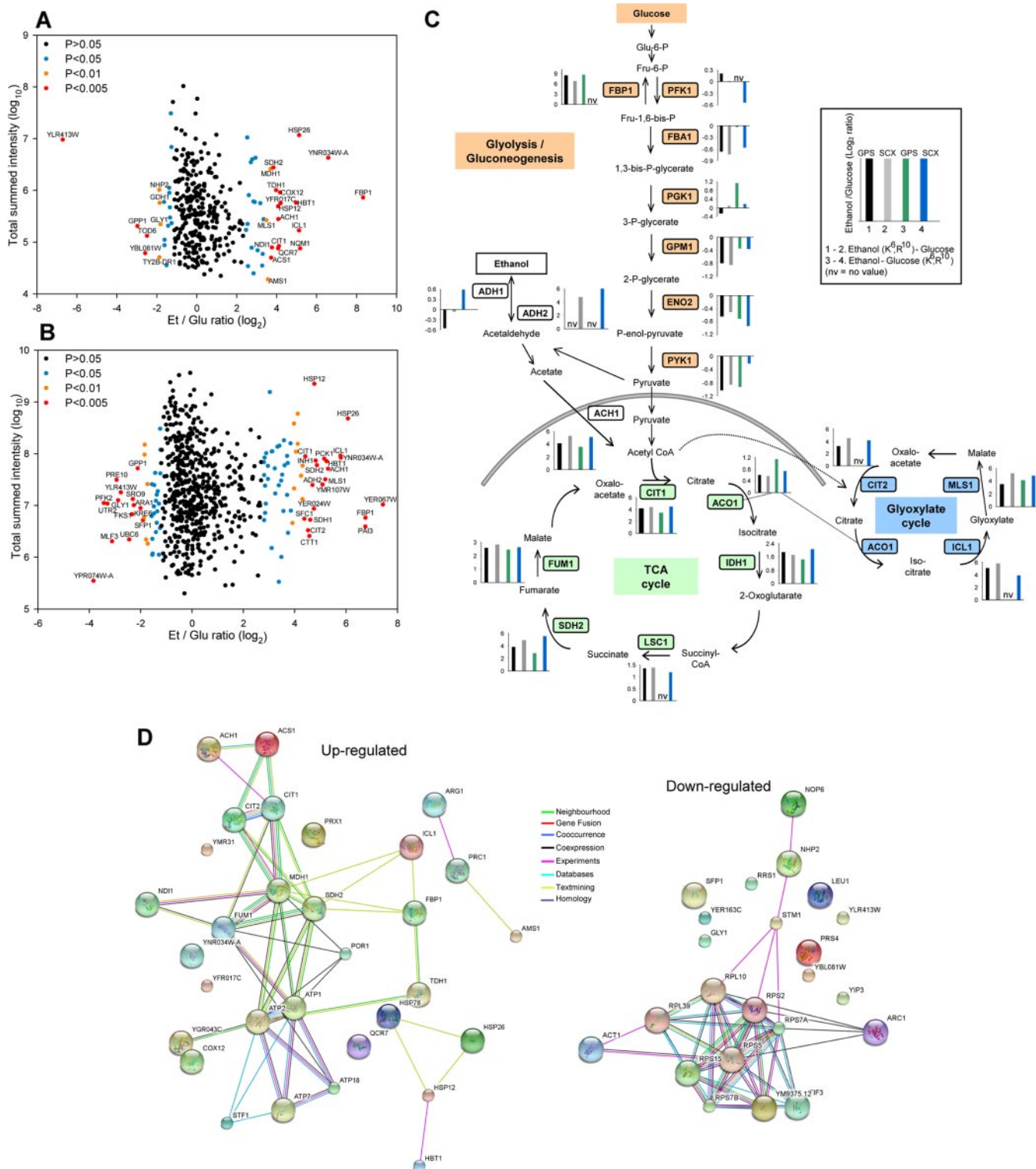
In this study, we have demonstrated the quantitative capability, reproducibility, and sensitivity of the GPS approach for proteomic profiling (discovery). In about 20 min, the crude proteome was reduced to a suitably large pool of motif-containing peptides, which was then subjected to LC-MS-MS. In comparison, reproducible detection and quantification of (low-abundant) proteins constitutes a major bottleneck in global shotgun profiling efforts using mass spectrometry (33–35). A common solution to this matter has been extensive conventional pre-fractionation prior to LC-MS-MS, but this approach, which might also introduce other issues (e.g. logistics, yield, and reproducibility), is not practical when large sample cohorts are to be addressed. In addition, the instrument time is frequently suboptimally used because of repetitive sampling of peptides from high-abundant proteins. In the case of GPS, the latter two issues could, if so required, be even bypassed, by (1) optimizing the capture step by designing the motifs so that few or no high-abundant motif-carrying

peptides are targeted, and (2) optimizing the assay step by multiplexing, *i.e.* using  $\geq 2$  CIMS antibodies at the same time (18, 22). In order to increase the throughput, MALDI mass spectrometry could also be adopted in the detection step depending on the complexity of the sample (21, 22).

The GPS was shown to display a dynamic range of at least three orders of magnitude, capable of targeting high- ( $>1$  million copies/cell) as well as low-abundant ( $<50$  copies/cell) proteins (Fig. 2), thereby extending our initial findings (22). Additional experiments will be required in order to determine in particular the quantitative dynamic range. Although the abundance levels, adopted from Ghememaghani *et al.* 2003 (32), were based on a different strain of *S. cerevisiae*, and the proteome was harvested slightly earlier ( $OD_{600}$  0.5 *versus* 0.7), a comparison of abundance levels was still relevant because the cells were harvested in log-phase in both cases. Moreover, the sensitivity (and throughput) could be further optimized by combining GPS with predefined isotopic labeled peptides present in the elution buffer and the multiple reaction monitoring (MRM) technique, paving the way for high-throughput proteomic discovery and validation efforts (19, 34–37).

The GPS technology was found to exhibit a peptide (protein) coverage overlapping and complementary to existing proteomic technologies (32), validating the findings reported in our initial study (22). As for example, in the spike-in yeast proteome experiments, 239 peptides, of which 53 (27%) were not previously reported in PeptideAtlas (38), were identified based on only two CIMS antibodies and a few LC-MS/MS runs (Fig. 1). The observation that GPS were found to target peptides of overlapping but different (shorter) average length than SCX could be explained by (1) the nature of the motif-containing peptides, (2) the affinity of the antibodies (assuming that shorter peptides binds more strongly), and/or (3) the





**FIG. 6. Biological relevance of identified proteins for carbon and ethanol metabolism.** The biological relevance of differentially expressed proteins in glucose *versus* ethanol yeast was evaluated. The GPS assay was based on six CIMS antibodies. **A**, all identified proteins by GPS for biosample 1 (ethanol ( $K^6, R^{10}$ ) - glucose) plotted as overall fold change as a function of protein intensity in the MS. Protein ratios are color-coded according to their ratio significance (significance B, as described in (29)). **B**, all identified proteins by SCX for biosample 1 (ethanol ( $K^6, R^{10}$ ) - glucose) plotted as overall fold change as a function of protein intensity in the MS. Protein ratios are color-coded according to their ratio significance (significance B, as described in (29)). **C**, GPS and SCX analysis of the central carbon metabolism. Protein expression values generated by GPS and SCX were mapped to the central carbon metabolism pathway(s) displayed in a condensed version according

nature of the SCX peptide population in the mass spectrometry detection step. Although the GPS approach identified about half the number of proteins as SCX (Fig. 3), it was difficult to draw any direct conclusions, because these numbers depend heavily on the frequency of the CIMS motifs, the affinity of the antibodies (normally in the sub- $\mu\text{M}$  range (22)), and the number of runs. In this case, the SCX set-up was assigned more LC-MS/MS time, and it is a well-known fact that the more MS/MS that are acquired the deeper one may probe the targeted proteome (33, 39).

On average, each CIMS antibody was found to target about 75 different motif-containing yeast proteins. In this context, it should be noted that the CIMS antibodies were originally selected against motifs designed for targeting the human proteome (22), clearly highlighting the cross-species applicability. Although only six of these non-yeast optimized CIMS antibodies were used along with minimal LC-MS/MS time, a significant section of the central carbon metabolism pathway in yeast was still covered using GPS (Figs. 6C and 6D). Notably, more than 60 proteins were uniquely identified with the GPS approach, again showing the complementary coverage provided compared with existing proteomic technologies, here illustrated by SCX. The coverage could be further extended, and directed against other pathways and/or certain protein families by generating additional CIMS binders against carefully designed CIMS-motifs. In this context, our recombinant scFv antibody library does not only represent a vast, renewable probe source, but could also provide unique opportunities in generating antibodies against peptides displaying low or even no immunogenicity (24). In addition, the proteome coverage could potentially also be extended by changing the specificity of the digesting enzyme employed and generating a new set of C-terminal motif specific CIMS antibodies. This would allow for a different set of peptides to be targeted, and it is a well-known fact that different portions of the proteome are more or less suited to be viewed by a mass spectrometry (40). Some peptides were detected and enriched by more than one CIMS antibody, but the cross talk was low. Some cross-binding will be anticipated, considering the nature of the experimentally determined binding motif, e.g. limited number of key residues (often only two or three), as well as nature and position of these residues. In addition, some of the overlapping peptides were identified as potential nonspecific background-binding peptides.

The quantitative capability and reproducibility was demonstrated addressing samples of different complexity (dynamic range). First, the quantitative accuracy was outlined by successfully determining the ratios ( $\leq$ fourfold changes) of pre-mixed non-SILAC and SILAC-labeled glucose grown yeast

proteomes (Figs. 1B and 1C). The observed groups of  $\log_2$  ratios were relatively skewed as compared with the expected ratios using non-normalized data, whereas the anticipated ratios were observed when the data was normalized toward the 50/50 ratio. Although no clear trend was observed, the data indicated that the accuracy of the ratio decreased with decreasing signal intensities. We chose to display the raw data intensities from MaxQuant in order to better reflect the true performances of GPS. However, in real comparative experiments, normalized data should be applied. Next, the technical reproducibility, *i.e.* including both capture and LC-MS/MS, was found to be high ( $R^2$ -values  $\leq 0.99$ ), whether proteomes with low ( $\leq$ fourfold changes) or high ( $\gg$ fourfold changes) were profiled, outlining the applicability of the GPS set-up for quantitative proteomics.

Furthermore, the determined expression ratios conformed very well with those determined for the same samples using SCX on the peptide level (Fig. 4). Noteworthy, this was the case also on the protein level, although the SCX data in several cases was based on multiple peptides per protein, whereas GPS mainly relied on a single peptide per protein, supporting the underlying concept of the GPS approach. In these comparisons, we used normalized data since the isotopically labeled conditions consistently were underrepresented in both GPS and SCX, although somewhat higher  $R^2$ -values were obtained for non-normalized data in both cases.

The poor reproducibility in terms of overlap of MS/MS identified peptides between two technical LC-MS/MS runs of complex peptide mixtures is well known and because of the stochastic nature of data dependent sampling (39). The GPS-generated fractions consistently outperformed the SCX fractions (Fig. 5) in terms of MS/MS identified peptides in separate LC-MS/MS runs. In more detail, the median identification reproducibility was 68% for GPS, but only 43% for SCX. Perhaps even more importantly, when comparing the overlap between biological replicate samples, the GPS methodology was found to display a median MS/MS identification reproducibility of 62% compared with only 31% for SCX. Notably, the SCX data (Fig. 5B) was based on the same fraction (eluate) injected twice for LC-MS/MS analysis (technical variations in the LC-MS/MS runs), whereas the GPS data (Fig. 5A) always was based on two separate captures and LC-MS/MS runs (*i.e.* technical variations for the entire assay). Noteworthy, when discussing the reproducibility, the peptide capture step was always included when appropriate, indicating that the variation introduced by the capture step had to be small. Hence, the data outlined the reproducibility of the GPS methodology and a key advantage for consistent and reproducible

---

to (34). D, Protein-protein interaction maps of GPS identified up- and down-regulated proteins for biosample 1. Thirty-one up-regulated proteins (limited to proteins reported/annotated with a single Swiss-Prot ID, MaxQuant significance B value of  $p < 0.05$  resulting in proteins with  $\geq 5.5$ -fold up-regulation). Twenty-three down-regulated proteins (limited to proteins reported/annotated with a single Swiss-Prot ID, MaxQuant significance B value of  $p < 0.1$  resulting in proteins with  $\geq 2$ -fold down-regulation).

measurement (profiling) of the same set of analytes (peptides) in various biological samples. In a recent protein expression study addressing technical and biological variation when applying ITRAQ and *S. cerevisiae* as model system, the highest variation was clearly attributed to the biological variation (41). In other words, adopting the GPS approach, running one capture per binder and biological sample might frequently be enough, thereby releasing valuable MS instrument time that instead could be used to run more biological replicate samples. In order to keep the GPS assay time at a minimum, multiplexing, *i.e.* applying several CIMS antibodies at the same time, will be essential.

Significant efforts have been made to map the proteome of yeast using various MS-based approaches (32, 34, 42). The present approach detected several differentially expressed proteins of high biological relevance in glucose *versus* ethanol cultivated yeast proteomes, representing the first GPS-based application (Fig. 6). In agreement with earlier findings using conventional proteomic methodologies (43, 44), many of these proteins were pin-pointed to the TCA-cycle and glyoxylate cycle, and displayed a significant induction upon growth in the presence of ethanol as carbon source. As for example, fructose-1,6-bisphosphatase I (FBP1), a key regulatory enzyme in the gluconeogenesis pathway (45, 46) was identified as being massively up-regulated. Furthermore, the expression trends for additional enzymes, such as MLS1, CIT1, ENO2, and SDH2 correlated well with a recent MRM study (47). In fact, only two key enzymatic players, including glucose-repressible alcohol dehydrogenase II (ADH2) and phosphoenolpyruvate carboxykinase 1 (PCK1) were missed by the GPS set-up, albeit using only six nonyeast optimized CIMS antibodies. It might also be of interest to note that two stress related proteins, HSP12 and HSP26; were highly up-regulated in the ethanol condition, a feature commonly observed for yeast cultivated under stressed conditions (48, 49). Furthermore, in agreement with other studies (50, 51), a set of several down-regulated ribosomal related proteins and related transcription, translation initiation factors (SFP1 and TIF3), and RNA-binding proteins (NOP6 and ARC1) were identified with the GPS setup. Of note, several uncharacterized proteins (YNR034W-A, YFR017C, YBL081W, YLR413W) were also identified to be significantly up- or down-regulated in both biological replicate sets using GPS. Further experiments will be required in order to delineate their functional role, and use in *e.g.* the development of ethanol tolerant yeast strains.

The GPS methodology thus currently display a promising suitability for addressing proteomes, such as mammalian tissue extracts and eukaryotes (*e.g.* yeast proteomes) in a discovery mode for identification and quantification of both high- and low-abundant analytes in a reproducible, cross-species manner using a limited set of CIMS binders. In other words, GPS provides a technology and an opportunity not matched by using either affinity proteomics or mass spectrometry (interfaced with classical fractionation methods) separately. The

approach is currently limited by i) the motif design, ii) antibody affinity *versus* proteome target, and iii) performance of the MS-MS setup. First, the motif design and subsequent antibody selection steps is still a key bottleneck if certain proteins or proteins groups are to be specifically targeted as well as avoided (*e.g.* high-abundant species). Second, the affinity of the CIMS antibodies (currently in the  $\mu\text{M}$  range) *versus* the target proteome might represent a limitation, because captured peptides could be lost, *i.e.* washed-off, when more complex proteomes, such as human serum, requiring even more stringent washing conditions in order to reduce nonspecific background binding, are targeted. Third, as for any MS-based approach, the performance of the selected MS instrumentation will be critical for the GPS set-up.

In conclusion, we have studied the applicability of the GPS platform (22), and demonstrated its quantitative capability, reproducibility, sensitivity, and coverage, outlining its potential within proteomic discovery profiling efforts. The GPS technology provides a novel methodological edge when addressing complex samples, and could be suitable for expression profiling studies, ranging from large-scale unbiased discovery studies to focused MRM type like assays, paving the way for the next generation of affinity proteomic efforts.

*Acknowledgments*—We thank Dr. Karin Hansson for assistance with the LTQ-Orbitrap instrument, Dr. Fredrik Levander, supported by Bioinformatics Infrastructure for Life Science (BILS), for all assistance with the Mascot server and Prof. Ole N. Jensen for kindly donating the yeast strain.

\* This work was supported by grants from Swedish National Research Council (VR-NT), the Foundation for Strategic Research (SSF) (Strategic Centre for Translational Cancer Research - CREATE Health), and Vinnova.

☐ This article contains supplemental Figs S1 to S6 and Tables S1 to S6.

§ To whom correspondence should be addressed: Department of Immunotechnology, Lund University, BMC D13, 221 84 Lund, Sweden. Tel.: +46-46-2224323; Fax: +46-46-2224200; E-mail: christer.wingren@immun.lth.se.

## REFERENCES

- Hanash, S. (2003) Disease proteomics. *Nature* **422**, 226–232
- Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**, 198–207
- Cravatt, B. F., Simon, G. M., and Yates, J. R., 3rd (2007) The biological impact of mass-spectrometry-based proteomics. *Nature* **450**, 991–1000
- Bengtsson, S., Krogh, M., Szigyarto, C. A., Uhlen, M., Schedvins, K., Silfverswärd, C., Linder, S., Auer, G., Alaiya, A., and James, P. (2007) Large-scale proteomics analysis of human ovarian cancer for biomarkers. *J. Proteome Res.* **6**, 1440–1450
- Whiteaker, J. R., Lin, C., Kennedy, J., Hou, L., Trute, M., Sokal, I., Yan, P., Schoenherr, R. M., Zhao, L., Voytovich, U. J., Kelly-Spratt, K. S., Krasnoselsky, A., Gafken, P. R., Hogan, J. M., Jones, L. A., Wang, P., Amon, L., Chodosh, L. A., Nelson, P. S., McIntosh, M. W., Kemp, C. J., and Paulovich, A. G. (2011) A targeted proteomics-based pipeline for verification of biomarkers in plasma. *Nat. Biotechnol.* **29**, 625–634
- Listgarten, J., and Emili, A. (2005) Practical proteomic biomarker discovery: taking a step back to leap forward. *Drug Discov. Today* **10**, 1697–1702
- Borrebaeck, C. A., and Wingren, C. (2007) High-throughput proteomics using antibody microarrays: an update. *Expert Rev. Mol. Diagn.* **7**, 673–686



8. Borrebaeck, C. A., and Wingren, C. (2009) Design of high-density antibody microarrays for disease proteomics: key technological issues. *J. Proteomics* **72**, 928–935
9. Haab, B. B. (2006) Applications of antibody array platforms. *Curr. Opin. Biotechnol.* **17**, 415–421
10. Kingsmore, S. F. (2006) Multiplexed protein measurement: technologies and applications of protein and antibody arrays. *Nat. Rev. Drug Discov.* **5**, 310–320
11. Ayoglu, B., Häggmark, A., Neiman, M., Igel, U., Uhlén, M., Schwenk, J. M., and Nilsson, P. (2011) Systematic antibody and antigen-based proteomic profiling with microarrays. *Expert Rev. Mol. Diagn.* **11**, 219–234
12. Carlsson, A., Wingren, C., Kristensson, M., Rose, C., Fernö, M., Olsson, H., Jernstrom, H., Ek, S., Gustavsson, E., Ingvar, C., Ohlsson, M., Peterson, C., and Borrebaeck, C. A. K. (2011) Molecular serum portraits in patients with primary breast cancer predict the development of distant metastases. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 14252–14257
13. Carlsson, A., Wuttge, D. M., Ingvarsson, J., Bengtsson, A. A., Sturfelt, G., Borrebaeck, C. A. K., and Wingren, C. (2011) Serum protein profiling of systemic lupus erythematosus and systemic sclerosis using recombinant antibody microarrays. *Mol. Cell. Proteomics* **10**, M110.005033
14. Sanchez-Carbayo, M., Socci, N. D., Lozano, J. J., Haab, B. B., and Cordon-Cardo, C. (2006) Profiling bladder cancer using targeted antibody arrays. *Am. J. Pathol.* **168**, 93–103
15. Dextrin-Mellby, L., Sandström, A., Centlow, M., Nygren, S., Hansson, S. R., Borrebaeck, C. A. K., and Wingren, C. (2010) Tissue proteome profiling of preeclamptic placenta using recombinant antibody microarrays. *Proteomics Clin. Appl.* **4**, 794–807
16. Borrebaeck, C. A., and Wingren, C. (2009) Transferring proteomic discoveries into clinical practice. *Expert Rev. Proteomics* **6**, 11–13
17. Poetz, O., Hoeppe, S., Templin, M. F., Stoll, D., and Joos, T. O. (2009) Proteome wide screening using peptide affinity capture. *Proteomics* **9**, 1518–1523
18. Wingren, C., James, P., and Borrebaeck, C. A. K. (2009) Strategy for surveying the proteome using affinity proteomics and mass spectrometry. *Proteomics* **9**, 1511–1517
19. Anderson, N. L., Anderson, N. G., Haines, L. R., Hardie, D. B., Olafson, R. W., and Pearson, T. W. (2004) Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *J. Proteome Res.* **3**, 235–244
20. Whiteaker, J. R., Zhao, L., Anderson, L., and Paulovich, A. G. (2010) An automated and multiplexed method for high throughput peptide immunoaffinity enrichment and multiple reaction monitoring mass spectrometry-based quantification of protein biomarkers. *Mol. Cell. Proteomics* **9**, 184–196
21. Hoeppe, S., Schreiber, T. D., Planatscher, H., Zell, A., Templin, M. F., Stoll, D., Joos, T. O., and Poetz, O. (2011) Targeting Peptide termini, a novel immunoaffinity approach to reduce complexity in mass spectrometric protein identification. *Mol. Cell. Proteomics* **10**, M110.002857
22. Olsson, N., Wingren, C., Mattsson, M., James, P., O'Connell, D., Nilsson, F., Cahill, D. J., and Borrebaeck, C. A. K. (2011) Proteomic analysis and discovery using affinity proteomics and mass spectrometry. *Mol. Cell. Proteomics* **10**, M110.003962
23. Wingren, C., and Borrebaeck, C. A. K. (2009) Antibody-based microarrays. *Methods Mol. Biol.* **509**, 57–84
24. Söderlind, E., Strandberg, L., Jirholt, P., Kobayashi, N., Alexeiva, V., Aberg, A. M., Nilsson, A., Jansson, B., Ohlin, M., Wingren, C., Danielsson, L., Carlsson, R., and Borrebaeck, C. A. (2000) Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nat. Biotechnol.* **18**, 852–856
25. Planatscher, H., Supper, J., Poetz, O., Stoll, D., Joos, T., Templin, M. F., and Zell, A. (2010) Optimal selection of epitopes for TXP-immunoaffinity mass spectrometry. *Algorithms for Mol. Biol.* **5**, 28
26. 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
27. Gruhler, A., Olsen, J. V., Mohammed, S., Mortensen, P., Faergeman, N. J., Mann, M., and Jensen, O. N. (2005) Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* **4**, 310–327
28. Grassl, J., Westbrook, J. A., Robinson, A., Borén, M., Dunn, M. J., and Clyne, R. K. (2009) Preserving the yeast proteome from sample degradation. *Proteomics* **9**, 4616–4626
29. 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
30. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190
31. Jensen, L. J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., Bork, P., and von Mering, C. (2009) STRING 8—a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.* **37**, D412–416
32. Ghaemmaghami, S., Huh, W. K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K., and Weissman, J. S. (2003) Global analysis of protein expression in yeast. *Nature* **425**, 737–741
33. de Godoy, L. M., Olsen, J. V., de Souza, G. A., Li, G., Mortensen, P., and Mann, M. (2006) Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system. *Genome Biol.* **7**, R50
34. Picotti, P., Bodenmiller, B., Mueller, L. N., Domon, B., and Aebersold, R. (2009) Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. *Cell* **138**, 795–806
35. Lange, V., Malmström, J. A., Didion, J., King, N. L., Johansson, B. P., Schäfer, J., Rameseder, J., Wong, C. H., Deutsch, E. W., Brusniak, M. Y., Bühlmann, P., Björck, L., Domon, B., and Aebersold, R. (2008) Targeted quantitative analysis of *Streptococcus pyogenes* virulence factors by multiple reaction monitoring. *Mol. Cell. Proteomics* **7**, 1489–1500
36. Addona, T. A., Abbatiello, S. E., Schilling, B., Skates, S. J., Mani, D. R., Bunk, D. M., Spiegelman, C. H., Zimmerman, L. J., Ham, A. J., Keshishian, H., Hall, S. C., Allen, S., Blackman, R. K., Borchers, C. H., Buck, C., Cardasis, H. L., Cusack, M. P., Dodder, N. G., Gibson, B. W., Held, J. M., Hiltke, T., Jackson, A., Johansen, E. B., Kinsinger, C. R., Li, J., Mesri, M., Neubert, T. A., Niles, R. K., Pulsipher, T. C., Ransohoff, D., Rodriguez, H., Rudnick, P. A., Smith, D., Tabb, D. L., Tegeler, T. J., Varyath, A. M., Vega-Montoto, L. J., Wahlander, A., Waldemason, S., Wang, M., Whiteaker, J. R., Zhao, L., Anderson, N. L., Fisher, S. J., Liebler, D. C., Paulovich, A. G., Regnier, F. E., Tempst, P., and Carr, S. A. (2009) Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nat. Biotechnol.* **27**, 633–641
37. Picotti, P., Rinner, O., Stallmach, R., Dautel, F., Farrah, T., Domon, B., Wenschuh, H., and Aebersold, R. (2010) High-throughput generation of reaction-monitoring assays for proteins and proteomes. *Nat. Methods* **7**, 43–46
38. Deutsch, E. W., Lam, H., and Aebersold, R. (2008) PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. *EMBO Rep* **9**, 429–434
39. Tabb, D. L., Vega-Montoto, L., Rudnick, P. A., Varyath, A. M., Ham, A. J., Bunk, D. M., Kilpatrick, L. E., Billheimer, D. D., Blackman, R. K., Cardasis, H. L., Carr, S. A., Clauser, K. R., Jaffe, J. D., Kowalski, K. A., Neubert, T. A., Regnier, F. E., Schilling, B., Tegeler, T. J., Wang, M., Wang, P., Whiteaker, J. R., Zimmerman, L. J., Fisher, S. J., Gibson, B. W., Kinsinger, C. R., Mesri, M., Rodriguez, H., Stein, S. E., Tempst, P., Paulovich, A. G., Liebler, D. C., and Spiegelman, C. (2010) Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **9**, 761–776
40. Swaney, D. L., Wenger, C. D., and Coon, J. J. Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *J. Proteome Res.* **9**, 1323–1329
41. Gan, C. S., Chong, P. K., Pham, T. K., and Wright, P. C. (2007) Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *J. Proteome Res.* **6**, 821–827
42. de Godoy, L. M., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Fröhlich, F., Walther, T. C., and Mann, M. (2008) Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455**, 1251–1254
43. Kolkman, A., Olsthoorn, M. M., Heeremans, C. E., Heck, A. J., and Slijper, M. (2005) Comparative proteome analysis of *Saccharomyces cerevisiae* grown in chemostat cultures limited for glucose or ethanol. *Mol. Cell. Proteomics* **4**, 1–11
44. Zaman, S., Lippman, S. I., Zhao, X., and Broach, J. R. (2008) How *Saccharomyces* responds to nutrients. *Annu. Rev. Genet.* **42**, 27–81
45. Hung, G. C., Brown, C. R., Wolfe, A. B., Liu, J., and Chiang, H. L. (2004)



- Degradation of the gluconeogenic enzymes fructose-1,6-bisphosphatase and malate dehydrogenase is mediated by distinct proteolytic pathways and signaling events. *J. Biol. Chem.* **279**, 49138–49150
46. Santt, O., Pfirrmann, T., Braun, B., Juretschke, J., Kimmig, P., Scheel, H., Hofmann, K., Thumm, M., and Wolf, D. H. (2008) The yeast GID complex, a novel ubiquitin ligase (E3) involved in the regulation of carbohydrate metabolism. *Mol. Biol. Cell* **19**, 3323–3333
47. Costenoble, R., Picotti, P., Reiter, L., Stallmach, R., Heinemann, M., Sauer, U., and Aebersold, R. (2011) Comprehensive quantitative analysis of central carbon and amino-acid metabolism in *Saccharomyces cerevisiae* under multiple conditions by targeted proteomics. *Mol. Syst. Biol.* **7**, 464
48. Lee, M. V., Topper, S. E., Hubler, S. L., Hose, J., Wenger, C. D., Coon, J. J., and Gasch, A. P. (2011) A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. *Mol. Syst. Biol.* **7**, 514
49. Alexandre, H., Ansanay-Galeote, V., Dequin, S., and Blondin, B. (2001) Global gene expression during short-term ethanol stress in *Saccharomyces cerevisiae*. *FEBS Lett.* **498**, 98–103
50. Pleiss, J. A., Whitworth, G. B., Bergkessel, M., and Guthrie, C. (2007) Rapid, transcript-specific changes in splicing in response to environmental stress. *Mol. Cell.* **27**, 928–937
51. Dinh, T. N., Nagahisa, K., Yoshikawa, K., Hirasawa, T., Furusawa, C., and Shimizu, H. (2009) Analysis of adaptation to high ethanol concentration in *Saccharomyces cerevisiae* using DNA microarray. *Bioprocess. Biosyst. Eng.* **32**, 681–688