MCP Papers in Press. Published on June 18, 2018 as Manuscript TIR117.000516

Stable isotope dynamic labeling of secretomes (SIDLS) identifies authentic secretory proteins released by cancer and stromal cells

Dean E. Hammond[1], J. Dinesh Kumar[1], Lorna Raymond[1], Deborah M. Simpson[2], Robert J. Beynon[2], Graham J. Dockray[1] and Andrea Varro[1].

[1]Department of Cellular and Molecular Physiology, Institute of Translational Medicine, University of Liverpool, Crown St, Liverpool, UK.

[2]Centre for Proteome Research, Institute of Integrative Biology, University of Liverpool, Crown St, Liverpool, UK.

Key words: cell secretion, secretome, SILAC, short term labeling, cancer biology, exocytosis

Short title: use of dynamic labeling to define secreted proteins

Address correspondence to:
Dr Dean Hammond
Department of Cell and Molecular Physiology
Institute of Translational Medicine
University of Liverpool
Crown Street
Liverpool L69 3BX
UK

Email: d.e.hammond@liverpool.ac.uk

Phone: +44 151 794 9566

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

Conflicts

The authors declare no conflicts

Contributions

DEH, performed experiments, analyzed data and wrote the manuscript; JDK, performed experiments, analyzed data and reviewed the manuscript; LR, performed experiments; DS, performed experiments; RJB, co-designed experiments, analyzed data and co-wrote the manuscript; GJD, co-designed experiments, contributed to data interpretation and co-wrote the manuscript; AV, conceived study, designed experiments, contributed to data interpretation and co-wrote the manuscript.

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

ABSTRACT

Analysis of secretomes critically underpins the capacity to understand the mechanisms determining interactions between cells and between cells and their environment. In the context of cancer cell micro-environments, the relevant interactions are recognised to be an important determinant of tumor progression. Global proteomic analyses of secretomes are often performed at a single time point and frequently identify both classical secreted proteins (possessing an N-terminal signal sequence), as well as many intracellular proteins, the release of which is of uncertain biological significance. Here, we describe a mass spectrometry-based method for stable isotope dynamic labeling of secretomes (SIDLS) that, by dynamic SILAC, discriminates the secretion kinetics of classical secretory proteins and intracellular proteins released from cancer and stromal cells in culture. SIDLS is a robust classifier of the different cellular origins of proteins within the secretome and should be broadly applicable to non-proliferating cells and cells grown in short term culture.

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

INTRODUCTION

Protein secretion critically supports a diverse range of cellular functions including cell-cell and cell-matrix interactions, as well as specialized functions such as hormone or digestive enzyme release. The constitutive secretion of proteins is a property of all cells, while regulated secretion (*i.e.* dependent on release of preformed stores after increased intracellular Ca^{2+}) occurs in specialized cells including neurons, endocrine and exocrine cells. It is now appreciated that an understanding of secretomes (the totality of secreted proteins) is of crucial importance in health and disease (1-4). For example, the secretomes of cancer and stromal cells contribute strongly to the cellular microenvironment that determines tumor progression (5). Thus, secretome studies have proven attractive both because they may provide insight into mechanisms of disease and because they facilitate the discovery of biomarkers that can be used for diagnosis, staging and monitoring of therapy.

Despite considerable progress in developing methods for secretome profiling (6-8) there remain problematical issues in interpretation of the data. Such studies frequently identify "classical" secreted proteins defined by an N-terminal signal sequence but they also identify many intracellular proteins, the apparent secretion of which is often of uncertain significance and not readily discriminated from tissue leakage/cell death (9). Interpretation is further compounded by the fact that many studies are performed at a single time-point, such that kinetic differences in the release of different components of the secretome are obscured. The classification of secretome proteins by gene ontology (GO) terms or predictions from computational tools/algorithms such as SignalP (10) or SecretomeP (11) can be used to segregate classically secreted proteins from intracellular proteins. However, experimental approaches that support this classification would be of obvious advantage. For example, a triple-labeling, single time-point approach was adopted by Kristensen and colleagues (12), in which they pointed out that the extent of labeling could be used to discriminate newly synthesized secretome proteins and those that were mobilized from pre-existing stores. Here,

we extend this thinking by describing a mass spectrometry (MS)-based strategy using stable isotope dynamic labeling of secretomes (SIDLS) that discriminates between classical secretory proteins and intracellular proteins within the secretome of cultured cells. The method differs from traditional SILAC, in which proteins are labeled for a fixed period to ensure all are fully labeled. Further, it differs from the single time-point pulsed SILAC approach (12) through dynamic labeling, in which the progressive incorporation of label into proteins is monitored over time. We demonstrate that a time dependency of labeling is of considerable value in the study of cell secretomes. A kinetic approach exploits the different labeling kinetics of classical secretory proteins that exhibit rapid incorporation of label compared with the much slower labeling of the bulk of intracellular proteins, even though some of the latter are present in the secretome. By monitoring the rate of incorporation of labeled amino acids into newly synthesized proteins as they appear in the media, we can differentiate those proteins that have been destined for secretion from those with low rates of labeling or low turnover relative to the growth rate of the cells, a feature of intracellular proteins.

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

EXPERIMENTAL PROCEDURES

Cell culture

Human primary cancer-associated myofibroblasts (CAMs) were derived from resected human esophageal squamous cancer tissue, obtained from patients as described previously (13). Esophageal squamous cell cancer cells, OE21, were purchased from American Type Culture Collection (Manassas, VA). All cells were maintained at 37°C, in 5% v/v CO_2, and cultured in DMEM, supplemented with 10% v/v FBS as previously described (14).

Stable isotopic dynamic labeling, mass spectrometry and protein identification

Cells (1 x 10^6) were seeded in complete medium (DMEM) in five T75 flasks giving 80% - 90% confluency, per flask. The following day, the cell-conditioned medium on each flask was changed to fresh 37°C heavy-labeled ([$^{13}C_6$]-labeled L-lysine) DMEM (10 mL volume per dish, serum-free). At the following time intervals – 30 min, 1 h, 2 h, 6 h and 24 h – all 10 mL of now heavy-labeled cell-conditioned DMEM from each flask was collected for subsequent secretome profiling (Fig. 1A), as follows. Each medium/secretome preparation was centrifuged at 800 x *g* for 7 min to remove debris and the protein component within each was concentrated by mixing, with agitation, with 25 µl StrataClean resin (Agilent Technologies Ltd., Wokingham, UK). The resin beads were washed twice in 25mM ammonium bicarbonate (ambic). Each secretome-loaded StrataClean suspension was re-suspended in 80 µL of 25 mM ambic and 5 µL of 1% (w/v) RapiGest (Waters, Hertfordshire, UK) in 25 mM ambic, prior to on-bead proteolytic digestion with trypsin (MS grade Trypsin Gold, Promega). The samples were heated at 80°C for 10 min after which proteins were reduced, by the addition of 5 µL of 60 mM DTT at 60°C for 10 min, before being cooled prior to addition of 5 µL of 180 mM iodoacetamide and incubation at RT for 30 min in the dark. Trypsin (1 µg) was added and the samples incubated at 37°C overnight on a rotary mixer. Peptide digests were subsequently acidified by the addition of 1 µL of trifluoroacetic acid (TFA) and incubated at 37°C for 45 min. Following centrifugation at 17,000 × *g* for 30 min, 10 µL of each clarified supernatant (peptide mixture) was prepared for nano LC-MS/MS. Peptide digests (2 µL) from each sample were

loaded onto a trap column (Acclaim PepMap 100, 2 cm × 75 µm inner diameter, C18, 3 µm, 100 Å) at 5 µl min^{-1} with an aqueous solution containing 0.1% (v/v) TFA and 2% (v/v) acetonitrile. After 3 min, the trap column was set in-line with an analytical column (Easy-Spray PepMap® RSLC 50 cm × 75 µm inner diameter, C18, 2 µm, 100 Å) (Dionex). Peptides were loaded in 0.1% (v/v) formic acid and eluted with a linear gradient of 3.8 – 40% buffer B (HPLC grade acetonitrile 80% (v/v) with 0.1% (v/v) formic acid) over 95 min at 300 nl min^{-1}, followed by a washing step (5 min at 99% solvent B) and an equilibration step (15 min at 3.8% solvent). All peptide separations were carried out using an Ultimate 3000 nano system (Dionex/Thermo Fisher Scientific). The column was operated at a constant temperature of 40°C and the LC system was coupled to a Q-Exactive mass spectrometer (Thermo Fisher), as described previously (15). The Q-Exactive was operated in data-dependent mode with survey scans acquired at a resolution of 70,000 at m/z 200. Up to the 10 most abundant peptides of charge state between 2+ and 4+ were selected for fragmentation by higher energy collisional dissociation with an isolation window of 2.0 Th and normalized collision energy of 30. The maximum ion injection times for the survey scan and the MS/MS scans were 250 and 100 ms, respectively, and the ion target value was set to 1E6 for survey scans and 1E5 for the MS/MS scans. Repetitive sequencing of peptides was minimized through dynamic exclusion of the sequenced peptides for 20 s.

Acquired MS data were searched and analyzed using Andromeda (16) and MaxQuant 1.5.8.3 (17) against a reviewed human UniProt protein database (date: 03/09/2016 containing 20,203 entries), using the default settings; briefly: the minimum required peptide length was seven amino acids long, trypsin/P was specified as the proteolytic enzyme and a single missed cleavage was allowed. Cysteine carbamidomethylation was set as a fixed modification and methionine oxidation was allowed as a variable modification. The initial precursor and fragment ion maximum mass deviations were set to 20 ppm and 0.5 Da, respectively. Peptide and protein false discovery rates (FDRs) were set to 1%, the ‘requant’ function activated and ‘match between runs’ enabled with the default parameters. In Supplemental material

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

(Supplemental Fig. 5), we include copies of relevant figures in the main text (Figs. 3A/B, 6A/B and 8A/B) where the requant function was disabled, to demonstrate that although the total number of proteins is lower, the results and conclusions of this study are unchanged.

Quantification and kinetic analysis of secretion

To analyze the rate of incorporation of heavy stable isotope-labeled amino acids into nascent proteins within secreted proteins, the MaxQuant peptide-level 'evidence.txt' output file was analyzed in detail. Initially, peptides from known contaminant proteins as well as those generated by proteolytic mis-cleavage events (thus potentially carrying >1 labeling site) were omitted. Although only lysine-terminated peptides have the potential to carry a dynamic stable isotope label for kinetic measurements, identification of secretome proteins was based on both arginine- and lysine-terminated tryptic peptide matches. It would, of course, be possible in the future to use both labeled lysine and arginine in an effort to increase the number of kinetically informative peptides but the principles of the method we describe here would not change as a consequence. Peptide mass spectral 'evidence data' for secretome proteins were then split into two lists, according to cell-line (OE21 or CAM).

For each peptide passing a 1% FDR threshold in the Andromeda search, the relative isotope abundance (RIA) was calculated at each time-point, if present in the MS data. RIA is expressed as abundance of heavy, labeled peptide (H), divided by the abundance of all (heavy + light) peptide (RIA_t = H/H+L). We applied a set of stringent criteria to produce high quality data-sets for each cell line analyzed. To model the secretome labeling trajectory, the RIA data for at least three time-points were used. We focused on peptides that had been identified and quantified, allowing RIA calculation, at more than one time-point in the labeling trajectory, so that we were effectively tracking their RIA behavior over time. Moreover, since the protein content of the secretome increases with time, we only analyzed peptide mass spectral data that included RIA data at 6 h and 24 h post exchange of culture medium. A small number of peptides were rejected from further analysis if they implied labeling profiles that could not be

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

biologically possible in this experimental system, specifically, where the calculated RIA at 30 min or 60 min after medium exchange was greater than that after 6 h or 24 h – these are likely artefacts. Because proteins at t=0 are completely unlabeled and for fully labeled proteins, RIA=1, we fitted a simplified version of the general first order equation:

$$RIAt = (1-exp(-k.t))$$

which generates the optimal fitted curve for a first order rise to plateau labeling (*k*) from an initial value of 0 to a final value of 1.0. Fitting was achieved using the nls() function in R.

To assess changes in the abundance of proteins identified from the 1% FDR Andromeda search in our secretomes, we summed the mass spectral peptide intensity reported by MaxQuant for labeled (heavy) and unlabeled (light) features to obtain a quantification value. These intensity values represent the summed eXtracted Ion Current (XIC) of all isotopic clusters associated with the identified peptide sequence. If more than one peptide was identified and quantified per protein, we calculated the mean abundance of labeled (heavy) and unlabeled (light) peptide species. This was used to monitor changes in abundance of each individual protein in the secretome with time which, for a physiologically secreted protein, should increase. To calculate a measure of the flux of each protein from intracellular to extracellular pools/compartments, we first measured the abundance [P] secreted over an 18 h period by subtracting the amount secreted after 6 h from that after 24 h. We then multiplied this with the first-order rate constant (*k*) at which newly synthesized protein acquired heavy isotopic label, to give flux (flux = k . [P]). In order to obtain protein-level kinetic data, RIA data for peptides belonging to the same protein were grouped together and fitted using the nls() function in R. Our high-quality data-sets for both cell lines were then cross-annotated with the output from SignalP (as described in Functional analysis, below), to explore the relationship(s) between labeling kinetics and predicted sub-cellular localisation. All mathematical modelling and data visualisations used R (v3.5.0) and ggplot2 (v2.0.1). Extraction and visualization of mass spectral isotopic patterns and XIC data were carried out using the 'RforProteomics'

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

package (1.15.0) (18). Abundance and kinetic plots are provided for every protein in Supplemental Material.

Functional analysis

All protein hits from the 1% FDR Andromeda search were subsequently used for subcellular localisation and GO enrichment analysis. The FASTA amino acid sequence for each protein identified in OE21 or CAM secretomes was extracted from the UniProt database and submitted to SignalP v4.1 (http://www.cbs.dtu.dk/services/SignalP/), to identify classically secreted proteins wherein a threshold SignalP D-score > 0.5 defines classical secretion, as described previously (10).

UniProt accession numbers for all proteins in each secretome collected over the labeling trajectory were loaded into the R/Bioconductor package 'clusterProfiler' (version 3.8.0 (19)) to allow GO over-representation analyses. We used the 'enrichGO' function together with 'compareCluster' to track changes in the functional enrichment profile with time, based on a hypergeometric distribution using a background list of all proteins in the *H. sapiens* annotation database. To remove redundant GO terms the 'simplify' function was applied using the "Wang" measure of semantic similarity (20) (similarity cut-off of 0.5), reporting only terms with the lowest FDR-adjusted p-values. A similar approach was taken to obtain GO functional profiles of our secretomes based on labeling kinetics.

Western blot analysis

Secretomes were probed by Western blot for selected classically secreted proteins, namely MMP1 (antibody BAF901, R&D Systems, Oxfordshire, UK), MMP3 (antibody AF913, R&D Systems), TGFB/TGFβig-h3 (antibody AF2935, R&D Systems) and SCG2 (antibody ab96589, Abcam, Cambridge, UK). In some experiments, the cells were pre-incubated for 30 min with 10 $\mu g.ml^{-1}$ brefeldin A (BFA; eBioscience, Ltd., Hatfield, UK), or 10 $\mu g.ml^{-1}$ cycloheximide (Sigma, Dorset, UK) and/or 1 μM ionomycin (Sigma-Aldrich, Poole, UK). Proteins were

resolved by sodium dodecyl sulfate–polyacrylamide gel electrophoresis and processed for Western blotting as described previously (13).

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

RESULTS

Experimental strategy

Secretomes, particulary in the early periods of incubation, are low abundance and we concentrated proteins by adsorption onto StrataClean, a silica-based bead preparation with a high affinity for protein. This captures all secretome proteins, and tryptic digests can be conducted directly on the beads – SDS-PAGE of bead eluate after digestion confirms completeness of digestion (results not shown). Peptides recovered from the on-bead digests were then used directly for LC-MS/MS. To establish the linearity of the StrataClean bead capture, we completed control experiments in which fresh 'virgin' (v) culture media was mixed in different proportions (specifically, 0:100, 20:80, 50:50, 80:20 and 100:0) with media that had been cell-conditioned (cc) for 24 h with CAM cells (Supplemental Fig. 1, Supplemental Table 1). StrataClean was used exactly as described in experimental procedures. The label free abundance of recovered proteins exhibited excellent linearity with protein load, confirming the quantitative performance of the protein capture method (Supplemental Fig. 1, Supplemental Table 1). Further, the assessment of the degree of labeling of protein captured by StrataClean is internally-controlled and thus independent of the quantity of protein.

Kinetic parameters for secretome proteins: abundance and flux

Each protein in a secretome has two kinetic parameters of relevance. The first is the change in abundance as a function of time. A protein that is actively secreted should accumulate in the medium, unless there is an opposing removal process that takes the protein back into the cell or which elicits extracellular degradation (a possibility, given the number of endopeptidases that are secreted from cells). Thus, abundance is not enough to define secretome kinetics. The second necessary parameter is flux, or the rate at which the protein flows from the intracellular pool to the extracellular space. Measurement of flux can also discriminate between classically secreted proteins and those that are released from the cell through necrotic or apoptotic changes, provided a labeling method is used to discriminate these pools. If a cell is supplied with labeled precursors, such as amino acids, the newly

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

synthesized and labeled *intracellular* proteins will enter and equilibrate with the unlabeled pre-existing pool. Thus, although newly synthesized proteins are fully labeled, they are diluted by a large, pre-existing pool of unlabeled protein, and thus the RIA is low. Subsequent leakage from the cell would reflect loss of this minimally-labeled mixture. If a small proportion of this pool is then released from the cell, the fraction of protein molecules that are labeled will be low. By contrast, proteins that are *classically secreted by the constitutive pathway* do not have a large intracellular reservoir to dilute the labeling of newly synthesized molecules, so all labeled proteins exiting the endoplasmic reticulum/Golgi are immediately secreted from the cell and will exhibit rapid acquisition of complete labeling in the medium. A third class of proteins, those of the *regulated secretory pathway* which is a feature of neural, exocrine and endocrine cells, can have a large, stored intracellular pool in secretory vesicles, and thus, newly synthesized protein should enter this pool and may exhibit relatively slow labeling. We reasoned that these different kinetic behaviors could be used to discriminate between classically secreted proteins and those derived from intracellular protein leakage.

Tensioned against protein turnover is the change in protein abundance. Secretome protein pools that expand would be expected to exhibit a rapid increase in label enrichment, consistent with a small intracellular pool and physiological secretion. By contrast, continued leakage of an intracellular protein, *e.g.* through cell damage, will reflect the extent of labeling of the intracellular pool, and unless this is an intrinsically high turnover protein, the degree of labeling will be low. To develop this logic further, a secreted protein pool that is static (not increasing) but which has a high degree of labeling must be subject to rapid removal from the extracellular pool (Fig. 1B). It follows that two measurements, changes in secretome protein abundance and the kinetic profile of labeling, could resolve proteins that are physiologically secreted from those that leak from the cell.

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

Rapid labeling and secretion in both cancer and stromal cells of a classically secreted protein but not an intracellular protein

In this study, a cancer (OE21) cell and a stromal cell (CAM) were labeled over 24 h with a stable isotope-labeled amino acid in the medium and the size of the secreted pool was assessed by summing the abundance of labeled (heavy, H) and unlabeled (light, L) peptide mass spectral features over time (0.5 to 24 h; Fig. 1C, D). Concurrently, the extent of labeling was assessed by monitoring changes in the relative isotope abundance (RIA, expressed as H/[H+L]) of these features, with time. Both cell types were monitored over time for incorporation of [$^{13}C_6$]lysine, and for protein abundance. To illustrate the concept: the peptide SQNPVQPIGPQTP***K*** from an established physiologically secreted protein, matrix metalloproteinase 1 (MMP1), could be detected in unlabeled form 30 min after incubating CAMs in [$^{13}C_6$]lysine-containing culture medium but, at this time, no labeled peptide was present (Fig. 1C, D). The unlabeled material defines MMP1 that was pre-synthesized prior to medium exchange and equilibration of the extracellular labeled amino acid with the intracellular tRNA pool. Over time, there was a progressive increase in labeled MMP1, highlighted by the shift in the relative peak heights of either XICs (Fig. 1C) or peptide precursor ion isotopic envelopes (Fig. 1D), as well as in change in RIA with time (RIA_t, Fig. 2A, red line). However, intracellular proteins, of which tubulin beta 3 (TBB3) is a representative example, incorporated virtually no label during the same period (Fig. 2A, blue line). Moreover, a comparison of the quantified abundance (summing both labeled and unlabeled peptides to obtain the total pool) revealed that whilst the output of MMP1 increased steadily over the period of the experiment, the abundance of TBB3 remained relatively static or even declined (Fig. 2B).

Secretome constituents meeting the criteria for kinetic analysis

We established a dataset of secretome proteins upon which a rigorous analysis of secretory kinetics might be made. After removing peptide spectrum matches (PSMs) from known contaminant proteins (included in the MaxQuant installation by default), we had a dataset with

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

92,480 PSMs: 48,933 in OE21 secretome samples corresponding to 2,109 unique proteins and 43,547 in CAM samples corresponding to 1,815 unique proteins. Approximately 9% of total PSMs (8,274) were generated by proteolytic mis-cleavage events and although those containing lysine residues could also report on labeling extent, they were excluded in this analysis. Arginine-terminating PSMs accounted for approximately 40% of the total (OE21, 19,243; CAM, 17,566) and were used for protein identification purposes but not for measurement of label incorporation. The remaining 55,671 PSMs (OE21, 29,690; CAM, 25,981) were lysine-terminated with a single instance of this amino acid; these mapped to 1,751 (OE21) and 1,484 (CAM) unique proteins. Further application of the stringent filtering criteria described in the experimental procedures section yielded 910 individual proteins from OE21 cells (RIA_t data generated from 13,585 PSMs) and 549 proteins from CAMs (RIA_t data generated from 11,313 PSMs). Pre- and post-filtered RIA_t datasets can be found in Supplemental Table 2.

Classically secreted proteins exhibit distinct secretory kinetics

For each protein in the reduced and filtered dataset, the labeling vector was measured as a first order rate constant (*k*) defining the rise to plateau labeling and, if appropriate, exchange into a pre-existing, unlabeled intracellular pool. Moreover, the quantified abundance (summing both labeled and unlabeled peptides to obtain the total pool) was used to determine the rate of pool expansion. These two parameters were then used to discriminate classically secreted proteins from intracellular proteins that were externalized by leakage. Complete labeling trajectory and abundance curves for individual proteins from both CAMs and OE21 cells are provided in the Supplemental material (Supplemental Fig. 2 [CAM] and Supplemental Fig. 3 [OE21]; red plots – labeling kinetics; blue plots – abundance). Many proteins were rapidly labeled and after a short period the RIA_t reached unity. By contrast, other proteins were barely labeled in the same time frame. Further exploration of the data indicated that well known classically secreted proteins (*e.g.* MMP1, MMP2, MMP3, TGFβig-h3) incorporated [$^{13}C_6$]lysine rapidly whilst known intracellular proteins (*e.g.* BAG3, COF1, SYSC, TBB5) were labeled

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

minimally. In addition, the abundance in the cell medium of classically secreted proteins increased while that of intracellular proteins remained either constant or declined (for abundance data-sets for both CAM and OE21 cells, see the blue plots in Supplemental Fig. 2 and Supplemental Fig. 3, respectively). These observations therefore support the hypothesis that the rate of labeling of a protein (*k*) is a reliable indicator of physiological secretion, *i.e.* proteins destined for secretion by virtue of an N-terminal signal sequence and with no large intracellular pool to dampen the incorporation of label. High *k* values are associated with those proteins that were operationally defined as secreted, whether based on direct knowledge or through bioinformatically-derived (SignalP) prediction (Fig 3A and 3B, see also Supplemental Fig. 5A and 5B with requant disabled). The SignalP D-cutoff score is a combined value from both signal-peptide and cleavage site prediction networks. The default thresholds resulting in a positive prediction of a signal peptide are 0.5 or 0.45 for eukaryotic proteins with or without a transmembrane domain, respectively (10). Based on our biology-driven SIDLS approach, these thresholds appear to be accurate (see grey dashed lines in Fig. 3A, CAMs; and 3B, OE21 cells).

Matching secretory kinetics and SignalP scores

To relate secretome kinetics to a protein classification we divided the secretome data into proteins with a SignalP score of >0.5 (classical secretory proteins) compared with those without (<0.5). From the secretomes of the two cell types it was readily possible to identify six different protein populations, namely classically secreted or not, those that were common to the two cells or those exclusive to OE21 or CAMs (Fig. 4). After 6 h of labeling, the RIA of CAM secretome proteins with SignalP >0.5 was clearly distinguishable from those with <0.5. This was also evident although less pronounced in OE21 cells at 24 h (Fig. 5). Rates of label exchange differed for individual proteins (Fig. 6A, CAMs; and Fig. 6B, OE21 cells) but the discrimination based on SignalP score was nonetheless impressive. There are proteins, however, that appear to acquire label very quickly, but which are not classically secreted proteins (*i.e.* with SignalP scores <0.5; for *e.g.*, see dashed lines in Fig. 6A and 6B, see also

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

Supplemental Fig. 5C and 5D with requant disabled). These initially raised concerns to us but their RIA_t data are suspicious in that they show an instant rise to plateau without any subsequent increase over time. Detailed inspection of the peptide-level chromatography and mass spectrometry data for these proteins revealed them to be erroneous measurements derived from mis-assignment of heavy/light peptide features by MaxQuant. In fact, in all these cases, there was no tandem MS evidence of *both* light and heavy features, and we believe these are anomalous data-points. The association between SignalP >0.5 and k >0.02 h^{-1} (combined k for SignalP >0.5 in OE21, Fig. 6B, red line) was highly significant for both CAMs and OE21 cells (Fisher exact text, $p<0.001$ in both cases). Combined fits of all proteins, in each cell type, sub-classified based on SignalP score (> or <0.5), indicated a >10-fold difference in the rate of label exchange (k) in CAMs (Fig. 6C), and an approximately 5-fold difference in OE21 cells (Fig. 6D).

A cut-off of $k \geq 0.1$ h^{-1} for the kinetics of a classically secreted protein was initially selected based on an empirical assessment of the data. Indeed, in support of this, inspection of Fig. 6C (red line) shows the predicted k for all proteins with SignalP >0.5 in CAMs was actually quite close to this cut off, *i.e.* 0.058 h^{-1}. However, the selection of the cut-off in k to differentiate classically secreted proteins may vary for different cell types depending on their biology. For example, there were lower rates of label exchange in the cancer cell line (OE21; Fig. 6D). Alternative approaches to defining the cut-off may prove useful in comparative studies between cell types (21). Taking a combined fit value of each secretome sub-classified by SignalP > or <0.5, which is valid also based on its relationship to k (Fig. 3), identified a distinct group of 108 / 549 (CAMs) and 57 / 910 (OE21) proteins that are 'true' physiologically secreted proteins.

Validation of the kinetics of classically secreted proteins

To determine whether the SIDLS kinetic profiles were compatible with those derived by other methods, we used orthogonal analysis by Western blots to define the kinetic responses of

three representative classically secreted proteins, MMP1, MMP3 and transforming growth factor-beta-induced protein ig-h3 (TGFBI / TGFβig-h3), following pharmacological treatments that arrest protein secretion. When translation was inhibited in CAMs with cycloheximide, the accumulation in the medium of MMP1 and TGFβig-h3 was already inhibited at 1 h and for these two as well as MMP2 was very clearly inhibited at 4 h (Fig. 7A, left). Similarly, using the ER-to-Golgi transport inhibitor, brefeldin A (BFA), to block progression along the secretory pathway, the appearance of MMP1 was suppressed at 1 h and for all three proteins at 2 h (Fig. 7A, right). The rate of appearance and abundance of all three proteins in the secretome measured using immunoblotting closely mirrors that determined by SIDLS (Fig. 7B).

Application of SIDLS to proteins exhibiting regulated secretion

Constitutively secreted proteins predominate in the secretomes of the two cells studied here. However, stromal myofibroblasts including CAMs may exhibit a modest regulated secretory pathway (22) and in this context it is interesting that a marker of the pathway, secretogranin-2 (SCG2), was identified and exhibited rapid label incorporation (k = 0.274 h^{-1}). Immunoblotting confirmed that secretion of SCG2 was stimulated by a short (30 min) stimulation with ionomycin, consistent with calcium-evoked exocytosis and the response was resistant to cycloheximide and BFA, consistent with release from storage vesicles (Fig. 7C). Thus, SIDLS can also be applicable to proteins of the regulated secretory pathway. The kinetics of labeling of these proteins in the medium may be slower than for constitutively secreted proteins although this may be offset to a substantial extent for those regulated secretory proteins that exhibit preferential secretion of newly synthesized material (23, 24).

Meta-analysis of classically secreted proteins in the CAM secretome

Classically secreted proteins included representatives of several important classes of secreted protein, notably extracellular matrix (ECM) proteins such as TGFB/TGFβig-h3, proteases (*e.g.* MMP1, MMP3, C1R), protease inhibitors (*e.g.* TIMP1, TIMP2, SERPINE1, SERPINE2),

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

chemokines (*e.g.* RARRES2), cytokines (*e.g.* the growth factors, VEGFC and CTGF) and growth factor-associated proteins (*e.g.* IGFBP3, -4, -5, -6), some of which have already been characterized in previous studies of myofibroblast secretomes (14, 22, 25). When flux is plotted against estimated *k*, this sub-set is readily distinguishable (Fig. 8, selected examples are highlighted, see also Supplemental Fig. 5E and 5F with requant disabled).

Using a SignalP D-score < or >0.5, in isolation, as a classifier of secretion (Fig. 9A [CAM] and Fig. 9C [OE21]), GO over-representation analysis revealed molecular function (GOMF) terms linked with binding to multiple ECM components for secreted proteins (SigP >0.5). Non-secreted proteins, however, showed an enrichment with terms associated primarily with translation and cell structure, consistent with the idea that these are simply 'leaked' proteins that function in these intracellular processes. Utilising our data-driven, SIDLS-determined classification of the secretome, in combination with SignalP (see Fig. 6C and 6D) to separate the secretome into a 'high *k*' group (*k* > combined fit for SignalP >0.5 proteins) and a 'low *k*' group (*k* < combined fit for SignalP >0.5 proteins) reveals an almost identical pattern of GOMF terms (Fig. 9B [CAM] and 9D [OE21]), adding further confidence to our SIDLS-mediated approach to secretome classification.

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

DISCUSSION

We describe a dynamic stable isotopic labeling- and mass spectrometry-based approach to characterise the physiological secretome of any cell that can be maintained in culture. Unlike traditional SILAC approaches, our Stable Isotope Dynamic Labeling of Secretomes (SIDLS) method exploits the kinetics of exchange from light-to-heavy stable isotopic labeling that occurs with protein synthesis *de novo* over relatively short labeling trajectories, thus dispensing with the need for exogenously-added serum factors required in more long-term cell cultures. This is beneficial as serum often contains an abundance of factors that influence the behavior and physiology of the culture system. SIDLS can confidently discriminate the secretion kinetics of physiologically-relevant classical secretory proteins from intracellular proteins that are released from cells either through damage during cell culture, apoptosis or 'leakage'. It is therefore a powerful classifier of the different cellular origins of proteins within the secretome and should be broadly applicable, allowing secretome characterisation of non-proliferating cells and cells only viable in short term culture. Embedding new knowledge of the rate of synthesis of the secretome constituents improves upon previously described approaches that rely on either the time-consuming labeling of cells to completion with heavy isotopic labels (traditional SILAC approaches), complex click chemistry approaches, or a combination of the two (12, 26-29).

Across two different cell lines we obtained global secretome identification of over 2,000 proteins and, in parallel, definitively determine the dynamic secretome behavior of a large proportion of these proteins, helping define their true intracellular origin and physiological role (see Supplemental Tables 2 and 3 for complete data-sets). As expected, classical secreted proteins show a shift in RIA from 0 towards 1 over time (as exemplified by MMP1 in Fig. 2A, red line). This is especially true in the non-transformed cells (CAMs), taking the SignalP score as a classifier of secretion or not (Fig. 5A). Our dynamic labeling strategy, combined with assessment of total abundance in the secretome, allowed us to distinguish, with high confidence, secreted protein assignments from erroneous measurements borne out of

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

chromatographic and MS errors (e.g. co-eluting peptide MS isotopic envelopes skewing RIA calculations), as logic dictates that the total abundance of any secreted protein in the secretome must increase with time. Several proteins that appear to be secreted readily, but which have no known extracellular function, fall into this bracket (*e.g.* RRP12 in OE21 cells; GRHL1 and AL9A1 in CAMs; CHMP3 in both OE21 cells and CAMs).

In our dataset, some classically secreted proteins show near identical behavior across both cell-lines (see Supplemental Fig. 4). Although small, this list of proteins identified in both CAM and OE21 secretomes shows that in general, a commonality exists in the secretome behavior of proteins between stromal and cancer cells that exist within the same microenvironment. But several proteins were removed from our data-sets during stringent filtering of the RIA_t data. Indeed, some other classical secreted proteins, for *e.g.* additional members of the MMP family, did not meet our stringent filtering criteria – for example, where tandem MS evidence existed for both light and heavy peptide features, but only at one time point in the labeling trajectory (all protein data is included in the raw data in Supplemental Tables 2 and 3). More in-depth proteomic analyses, for *e.g.* adopting fractionation approaches of each secretome sample, would increase the number of proteins identified allowing improved cross-comparison(s) to be made. However, it must be noted that, in general, the relationship between the rate of labeling in OE21 cancer cells and rate of labeling in CAMs, for common proteins, is not a strong correlation. Much clearer is a generally lower rate of incorporation of label into newly synthesized protein in the cancer cells (OE21), indicative of defective protein synthesis and/or trafficking though the secretory system in cancer.

One of the main advantages of SIDLS is that it provides an orthogonal perspective to secretome dynamics. By tracking the appearance of label in secreted proteins, it is possible to build a profile of the speed and duration of response of individual proteins and resolve true secreted proteins from low level intracellular leakage. The marked consonance between proteins that would be labeled as secreted through a high predictive score of a signal peptide

and rapid labeling gives a convincing confirmatory perspective on the secretome. For this analysis, we have been very stringent in the retention of proteins, and those for which abundances were too low for recovery by data-dependent acquisition approaches would be recovered by more targeted methods, such as selected reaction monitoring (30). It is not too bold to imagine that the use of different labeled precursors in a pulse-labeling strategy would provide new insights into the interaction of co-cultures of cells mediated by their secretomes. Thus, inclusion of the simple expedient of dynamic labeling of secretomes will greatly increase the confidence with which such secretomes are studied.

ACKNOWLEDGMENTS

Supported by a grant from North West Cancer Research. We are grateful for support from the University of Liverpool Technology Directorate and the exceptional instrument support provided by Dr Philip Brownridge.

DATA AVAILABILITY

The mass spectrometry proteomics data for the SIDLS dynamic labeling aspect of this study have been deposited to the ProteomeXchange Consortium via the PRIDE (31) partner repository with the dataset identifier PXD007231. Equivalent data for the linearity of protein capture by StrataClean are deposited to ProteomeXchange too, with the identifier PXD009838. The output of the MaxQuant searches of the SIDLS dynamic labeling proteomics data are available using MS-Viewer (http://msviewer.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer), using the search key identifier: 88dyh3qzvc . All annotated spectra can be accessed here. That for the StrataClean linearity experiment is also available using MS-Viewer using the search key idenitifier: hjhr9jxzpk.

References

1. Ranganath, S. H., Levy, O., Inamdar, M. S., and Karp, J. M. (2012) Harnessing the mesenchymal stem cell secretome for the treatment of cardiovascular disease. *Cell Stem Cell* 10, 244-258

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

2. Alvarez-Llamas, G., Szalowska, E., de Vries, M. P., Weening, D., Landman, K., Hoek, A., Wolffenbuttel, B. H., Roelofsen, H., and Vonk, R. J. (2007) Characterization of the human visceral adipose tissue secretome. *Mol Cell Proteomics* 6, 589-600
3. Makridakis, M., and Vlahou, A. (2010) Secretome proteomics for discovery of cancer biomarkers. *J Proteomics* 73, 2291-2305
4. Wu, C. C., Hsu, C. W., Chen, C. D., Yu, C. J., Chang, K. P., Tai, D. I., Liu, H. P., Su, W. H., Chang, Y. S., and Yu, J. S. (2010) Candidate serological biomarkers for cancer identified from the secretomes of 23 cancer cell lines and the human protein atlas. *Mol Cell Proteomics* 9, 1100-1117
5. Hanahan, D., and Weinberg, R. A. (2011) Hallmarks of cancer: the next generation. *Cell* 144, 646-674
6. Holmberg, C., Ghesquiere, B., Impens, F., Gevaert, K., Kumar, J. D., Cash, N., Kandola, S., Hegyi, P., Wang, T. C., Dockray, G. J., and Varro, A. (2013) Mapping proteolytic processing in the secretome of gastric cancer-associated myofibroblasts reveals activation of MMP-1, MMP-2, and MMP-3. *J Proteome Res* 12, 3413-3422
7. Rieckmann, J. C., Geiger, R., Hornburg, D., Wolf, T., Kveler, K., Jarrossay, D., Sallusto, F., Shen-Orr, S. S., Lanzavecchia, A., Mann, M., and Meissner, F. (2017) Social network architecture of human immune cells unveiled by quantitative proteomics. *Nat Immunol* 18, 583-593
8. Gauthier, N. P., Soufi, B., Walkowicz, W. E., Pedicord, V. A., Mavrakis, K. J., Macek, B., Gin, D. Y., Sander, C., and Miller, M. L. (2013) Cell-selective labeling using amino acid precursors for proteomic studies of multicellular environments. *Nat Methods* 10, 768-773
9. Brown, K. J., Formolo, C. A., Seol, H., Marathi, R. L., Duguez, S., An, E., Pillai, D., Nazarian, J., Rood, B. R., and Hathout, Y. (2012) Advances in the proteomic investigation of the cell secretome. *Expert Rev Proteomics* 9, 337-345
10. Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8, 785-786
11. Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G., and Brunak, S. (2004) Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng Des Sel* 17, 349-356
12. Kristensen, L. P., Chen, L., Nielsen, M. O., Qanie, D. W., Kratchmarova, I., Kassem, M., and Andersen, J. S. (2012) Temporal profiling and pulsed SILAC labeling identify novel secreted proteins during ex vivo osteoblast differentiation of human stromal stem cells. *Mol Cell Proteomics* 11, 989-1007
13. McCaig, C., Duval, C., Hemers, E., Steele, I., Pritchard, D. M., Przemeck, S., Dimaline, R., Ahmed, S., Bodger, K., Kerrigan, D. D., Wang, T. C., Dockray, G. J., and Varro, A. (2006) The role of matrix metalloproteinase-7 in redefining the gastric microenvironment in response to Helicobacter pylori. *Gastroenterology* 130, 1754-1763
14. Kumar, J. D., Holmberg, C., Kandola, S., Steele, I., Hegyi, P., Tiszlavicz, L., Jenkins, R., Beynon, R. J., Peeney, D., Giger, O. T., Alqahtani, A., Wang, T. C., Charvat, T. T., Penfold, M., Dockray, G. J., and Varro, A. (2014) Increased expression of chemerin in squamous esophageal cancer myofibroblasts and role in recruitment of mesenchymal stromal cells. *PLoS One* 9, e104877
15. 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
16. 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
17. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* 26, 1367-1372
18. Gatto, L., Breckels, L. M., Naake, T., and Gibb, S. (2015) Visualization of proteomics data using R and bioconductor. *Proteomics* 15, 1375-1389

19. Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284-287
20. Yu, G., Li, F., Qin, Y., Bo, X., Wu, Y., and Wang, S. (2010) GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* 26, 976-978
21. Budczies, J., Klauschen, F., Sinn, B. V., Gyorffy, B., Schmitt, W. D., Darb-Esfahani, S., and Denkert, C. (2012) Cutoff Finder: a comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization. *PLoS One* 7, e51862
22. Balabanova, S., Holmberg, C., Steele, I., Ebrahimi, B., Rainbow, L., Burdyga, T., McCaig, C., Tiszlavicz, L., Lertkowit, N., Giger, O. T., Oliver, S., Prior, I., Dimaline, R., Simpson, D., Beynon, R., Hegyi, P., Wang, T. C., Dockray, G. J., and Varro, A. (2014) The neuroendocrine phenotype of gastric myofibroblasts and its loss with cancer progression. *Carcinogenesis* 35, 1798-1806
23. Varro, A., Dockray, G. J., Bate, G. W., Vaillant, C., Higham, A., Armitage, E., and Thompson, D. G. (1997) Gastrin biosynthesis in the antrum of patients with pernicious anemia. *Gastroenterology* 112, 733-741
24. Rhodes, C. J., and Halban, P. A. (1987) Newly synthesized proinsulin/insulin and stored insulin are released from pancreatic B cells predominantly via a regulated, rather than a constitutive, pathway. *J Cell Biol* 105, 145-153
25. Holmberg, C., Quante, M., Steele, I., Kumar, J. D., Balabanova, S., Duval, C., Czepan, M., Rakonczay, Z., Jr., Tiszlavicz, L., Nemeth, I., Lazar, G., Simonka, Z., Jenkins, R., Hegyi, P., Wang, T. C., Dockray, G. J., and Varro, A. (2012) Release of TGFbetaig-h3 by gastric myofibroblasts slows tumor growth and is decreased with cancer progression. *Carcinogenesis* 33, 1553-1562
26. Eichelbaum, K., and Krijgsveld, J. (2014) Combining pulsed SILAC labeling and click-chemistry for quantitative secretome analysis. *Methods Mol Biol* 1174, 101-114
27. Henningsen, J., Blagoev, B., and Kratchmarova, I. (2014) Analysis of secreted proteins using SILAC. *Methods Mol Biol* 1188, 313-326
28. Rocha, B., Calamia, V., Blanco, F. J., and Ruiz-Romero, C. (2016) Identification of Factors Produced and Secreted by Mesenchymal Stromal Cells with the SILAC Method. *Methods Mol Biol* 1416, 551-565
29. Roelofsen, H., Dijkstra, M., Weening, D., de Vries, M. P., Hoek, A., and Vonk, R. J. (2009) Comparison of isotope-labeled amino acid incorporation rates (CILAIR) provides a quantitative method to study tissue secretomes. *Mol Cell Proteomics* 8, 316-324
30. Holman, S. W., Hammond, D. E., Simpson, D. M., Waters, J., Hurst, J. L., and Beynon, R. J. (2016) Protein turnover measurement using selected reaction monitoring-mass spectrometry (SRM-MS). *Philos Trans A Math Phys Eng Sci* 374
31. Vizcaino, J. A., Csordas, A., Del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q. W., Wang, R., and Hermjakob, H. (2016) 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 44, 11033

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

FIGURE LEGENDS

Fig. 1. Concept, design and analytical structure of the study.

Panel *A*, Dishes of cells (OE21 or CAMs) were initially seeded and established in standard growth medium (DMEM) before this was exchanged for SILAC DMEM containing heavy stable isotope-labeled lysine ([$^{13}C_6$]lysine). Cell-conditioned SILAC DMEM was then collected at the indicated times (1 dish per time-point) for up to 24 h. Protein secreted from the cells was concentrated using StrataClean and digested into peptides for subsequent analysis by LC-MSMS. Panel *B*, Schematic diagram showing how measuring changes in the abundance of secreted protein(s) and their extent of labeling with heavy stable isotopic amino acids over time will resolve proteins that are physiologically secreted from those that leak from the cell. The rate of incorporation of [$^{13}C_6$]lysine into nascent proteins within the secretome was determined from the LC-MSMS data; *C*, shows representative extracted ion currents (XICs) and *D*, isotopic envelopes for the peptide SQNPVQPIGPQTPK from matrix metalloproteinase 1 (MMP1). Measuring the rate of incorporation of heavy isotope (expressed as the relative isotope abundance, or RIA), with time (RIA_t), allows secretion dynamics to be determine for every detected protein secreted.

Fig. 2. Representative dynamic behavior of a classically secreted protein (MMP1) and an intracellular protein (TBB3).

Panel *A*, Secreted proteins acquire heavy label quickly since there is no intracellular pool to dilute/delay acquisition of label. This is exemplified by the rapid rise towards RIA of 1, as shown by MMP1 (red line). Panel *B*, The abundance of a secreted protein should increase continuously with time, as also demonstrated by MMP1 (red line). Opposing behaviors can be seen for the intracellular protein TBB3 (*A*, *B*, blue lines).

Fig. 3. Rates of label exchange for individual proteins discriminated on the basis of SignalP score.

The first order rate constant (*k*) of incorporation of [$^{13}C_6$]lysine into nascent proteins within the CAM (*A*) or OE21 (*B*) secretome was plotted as a function of the SignalP score. Symbols are color-coded on the basis of SignalP score, from 0 = dark blue, through to the maximum score of 1 = red. SIDLS-determined predictions of secretion on the basis of rate (*k*) of light-to-heavy label incorporation closely match those predicted by SignalP, with the vast majority of proteins showing high rates of label exchange (> 0.05) also having high SignalP scores, typically >0.5. An equivalent plot with the MaxQuant requant function disabled is provided in Supplemental Material.

Fig. 4. Classification of proteins on the basis of secretome behavior.

Monitoring the abundance of each protein in the 24 h secretomes of the two cell types (OE21 and CAM) allows the identification of six different protein populations; classically secreted (SignalP score >0.5, large red points) or not (SignalP score <0.5, small grey points), and either common to the two cells (blue dashed circled) or exclusive to OE21 or to CAMs, as highlighted on the plot.

Fig. 5. Time-dependent changes in peptide relative isotope abundance (RIA).

For all high quality [$^{13}C_6$]lysine-labeled peptides at the indicated times, the RIA, expressed as the ratio H/(H+L) was calculated and plotted as a distribution curve using kernel density estimation. For proteins that do not contain a signal peptide (SignalP <0.5), peptides remain almost entirely unlabeled over the total trajectory of the study (blue lines). Peptides from secreted proteins (SignalP >0.5), however, demonstrate a clear transition from largely unlabeled to extensively labeled as a consequence of protein turnover/label exchange. This is especially true in the stromal CAM cell line (left hand panels), but is also evident although less pronounced in cancer (OE21) cells. The dotted lines define the maximum possible RIA that the system can attain.

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

Fig. 6. Rates of label exchange for individual proteins and complete secretomes, discriminated on the basis of SignalP scores.

First-order rate constants at which newly synthesized proteins acquire heavy isotopic label, for every protein in the CAM (Panel *A*) and OE21 (Panel *B*) secretomes. Physiologically-secreted proteins (SignalP >0.5; red lines) clearly acquire new [$^{13}C_6$]lysine at a higher rate compared to intracellular proteins that merely 'leak' from the cell (SignalP <0.5; blue lines). Some proteins with low SignalP scores appear to be readily secreted or have very high turnover (blue dashed lines); however, manual inspection of the raw MS data for these proteins revealed them to be artefacts (see main text for explanation). Relating 'global' secretome kinetics to protein classification on the basis of a computational prediction of the presence or not of a signal peptide (SignalP score), reveals an impressive discrimination in CAMs (Panel *C*) and, to a lesser extent, in cancer cells (Panel *D*), thus raising questions about the secretome behavior of cells *in vivo* in the tumor microenvironment. An equivalent plot with the MaxQuant requant function disabled is provided in Supplemental Material.

Fig. 7. Validation of the kinetics of classically secreted proteins by Western blotting.

Panel *A*, The kinetic behavior of three representative classically secreted proteins, MMP1, MMP3 and TGFBI/TGFβig-h3, following pharmacological blockade of secretion in CAMs, were validated using a standard immunoblotting approach. When translation was inhibited with cycloheximide, the accumulation in the medium of MMP1 and TGFβig-h3 was already reduced at 1 h. For these two proteins (and MMP2), secretion was clearly inhibited at 4 h. Similarly, upon perturbation of secretion using brefeldin A (BFA), the appearance of MMP1 was suppressed at 1 h and of all three proteins from 2 h. Panel *B*, The kinetics of secretion of MMP1, MMP3 and TGFβig-h3 determined by SIDLS clearly match those determined by the orthogonal approach of Western blotting. Panel *C*, Evidence that SIDLS can discern proteins of the regulated secretory pathway. Western blot confirmation that secretion of secretogranin-2 (SCG2), an established marker of a minor, endocrine-like secretory phenotype known to be

present in stromal myofibroblasts (CAMs), was stimulated by a short (30 min) stimulation with ionomycin. This is consistent with calcium-evoked exocytosis and the response was resistant to cycloheximide (left) and BFA (right), supporting the idea that this represents release from intracellular storage vesicles.

Fig. 8. Relationship between *k* and flux of protein from intracellular to extracellular pools.

A surrogate measurement of flux into the extracellular protein pool was determined by calculating, flux = *k* . [P], where *k* is first-order rate constant at which newly synthesized protein acquired heavy isotopic label and P is the amount of protein secreted over an 18 h period. Symbols are color-coded on the basis of SignalP classification (>0.5 = red; <0.5 = blue) with alpha transparency shading according to the score from SignalP predictions (from 0, high transparency – to 1, low transparency). Panel *A*, CAMs; Panel *B*, OE21 cells. Selected proteins of interest in tumor biology are highlighted. An equivalent plot with the MaxQuant requant function disabled is provided in Supplemental Material.

Fig. 9. Functional analysis of the secretome.

GO enrichment analysis was performed using the 'compareCluster' function in the R/Bioconductor package 'clusterProfiler'. All proteins identified in the CAM (Panel *A*, *B*) or OE21 (Panel *C*, *D*) secretomes were loaded simultaneously in clusters based upon either their classification according to SignalP (D-score < or >0.5, Panel *A*, *C*) or their classification according to SIDLS ('high *k*' vs 'low *k*', Panel *B*, *D*) and statistically over-represented GOMF terms of each protein set determined (hypergeometric test pvalueCutoff, p = 0.000001 for CAM, p = 0.01 for OE21). To aid visualisation and interpretation of the data, redundant GO terms were removed using the 'simplify' function (as described in Experimental Procedures) and the most over-represented GOMF terms plotted as dotplots. GO terms linked with binding

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

to multiple ECM components are over-represented in the secreted protein clusters (SigP >0.5 and 'high *k*'). Non secreted proteins, however, showed an enrichment with terms associated primarily with translation and cell structure (SigP <0.5 and 'low *k*'). Abbreviations: bdg. = binding; struct. = structural; constit. = constituent; act. = activity; transl. = translation; dep. = dependent; mol. = molecular.

ASBMB MOLECULAR & CELLULAR PROTEOMICS MCP

ASBMB
MOLECULAR & CELLULAR PROTEOMICS
MCP

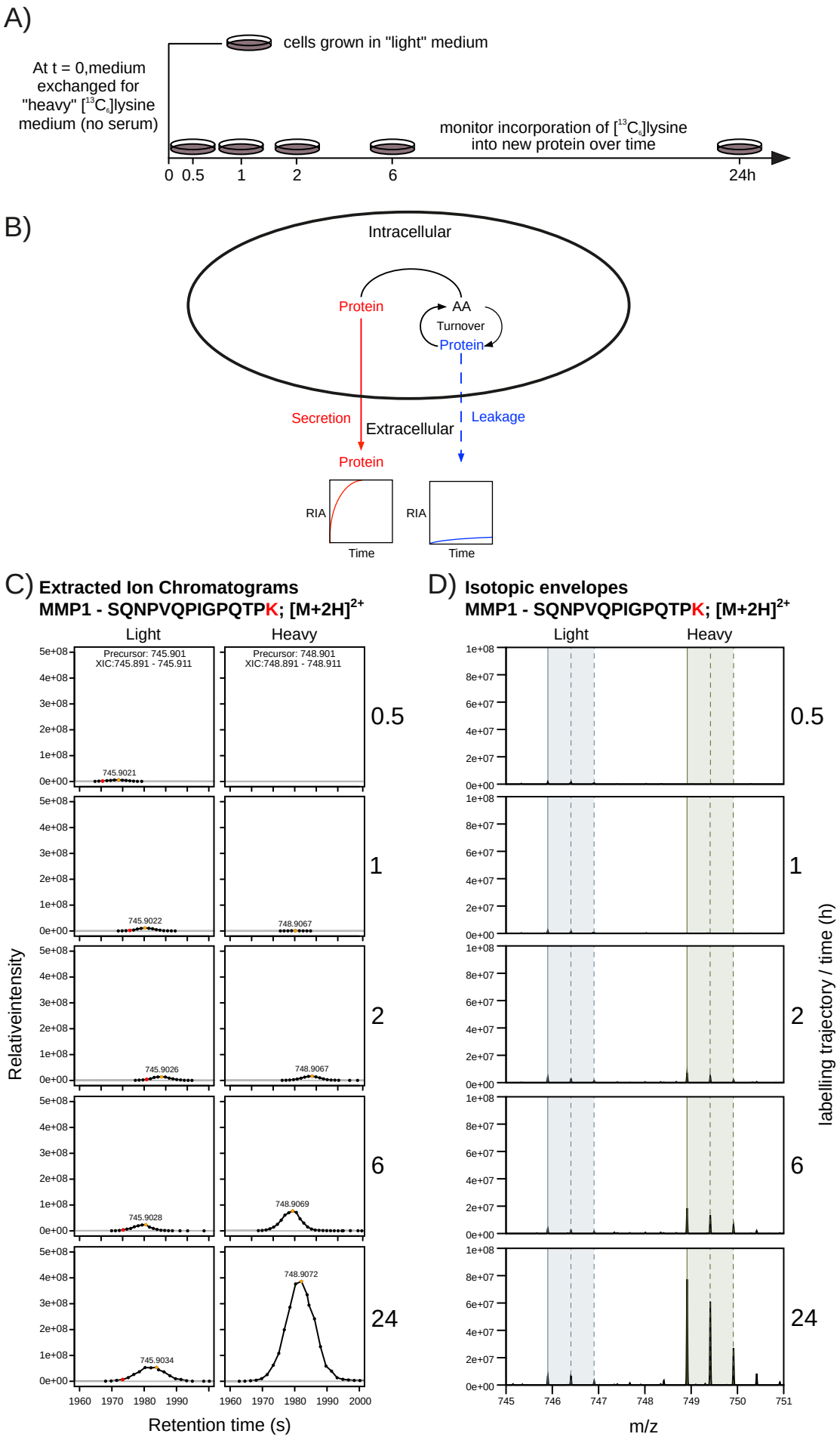
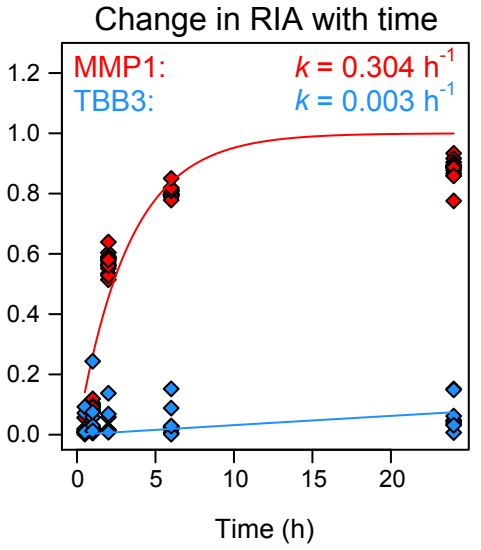


Figure 1

ASBMB
MOLECULAR & CELLULAR PROTEOMICS
MCP



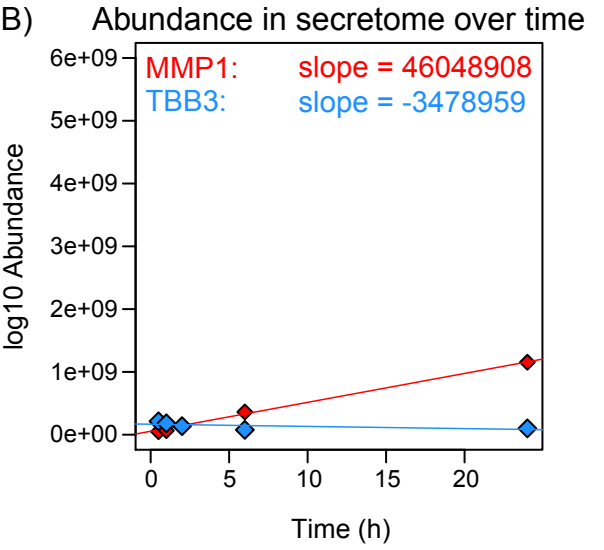


Figure 2

ASBMB
MOLECULAR & CELLULAR PROTEOMICS
MCP

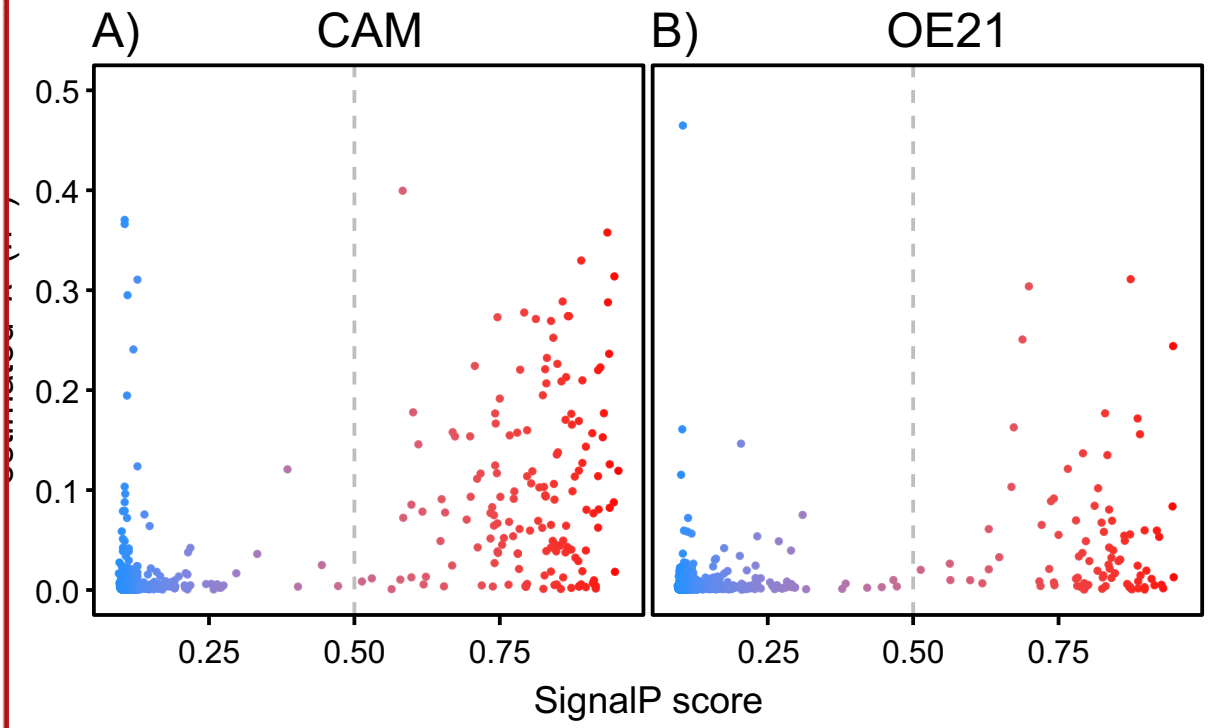


Figure 3

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

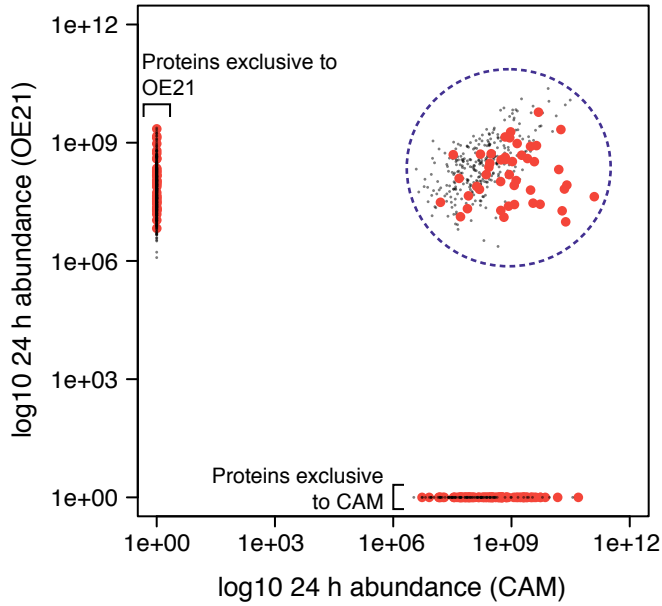


Figure 4

ASBMB
MOLECULAR & CELLULAR PROTEOMICS
MCP

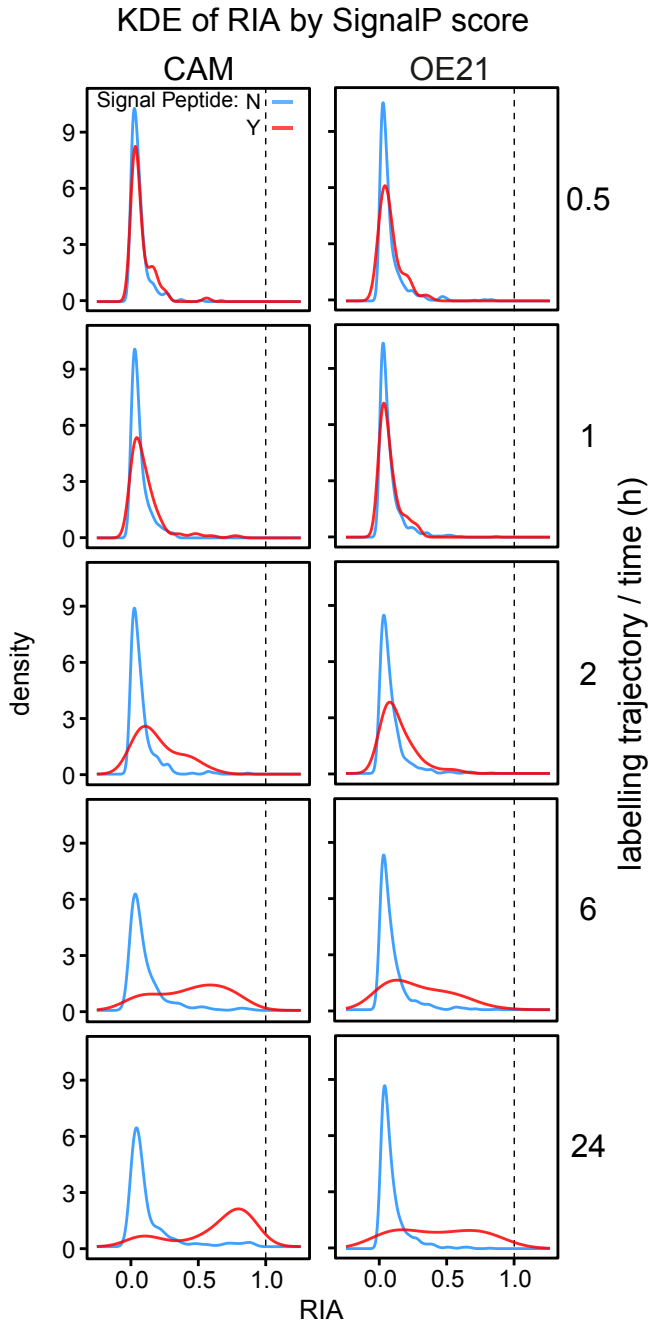


Figure 5

ASBMB
MOLECULAR & CELLULAR PROTEOMICS
MCP

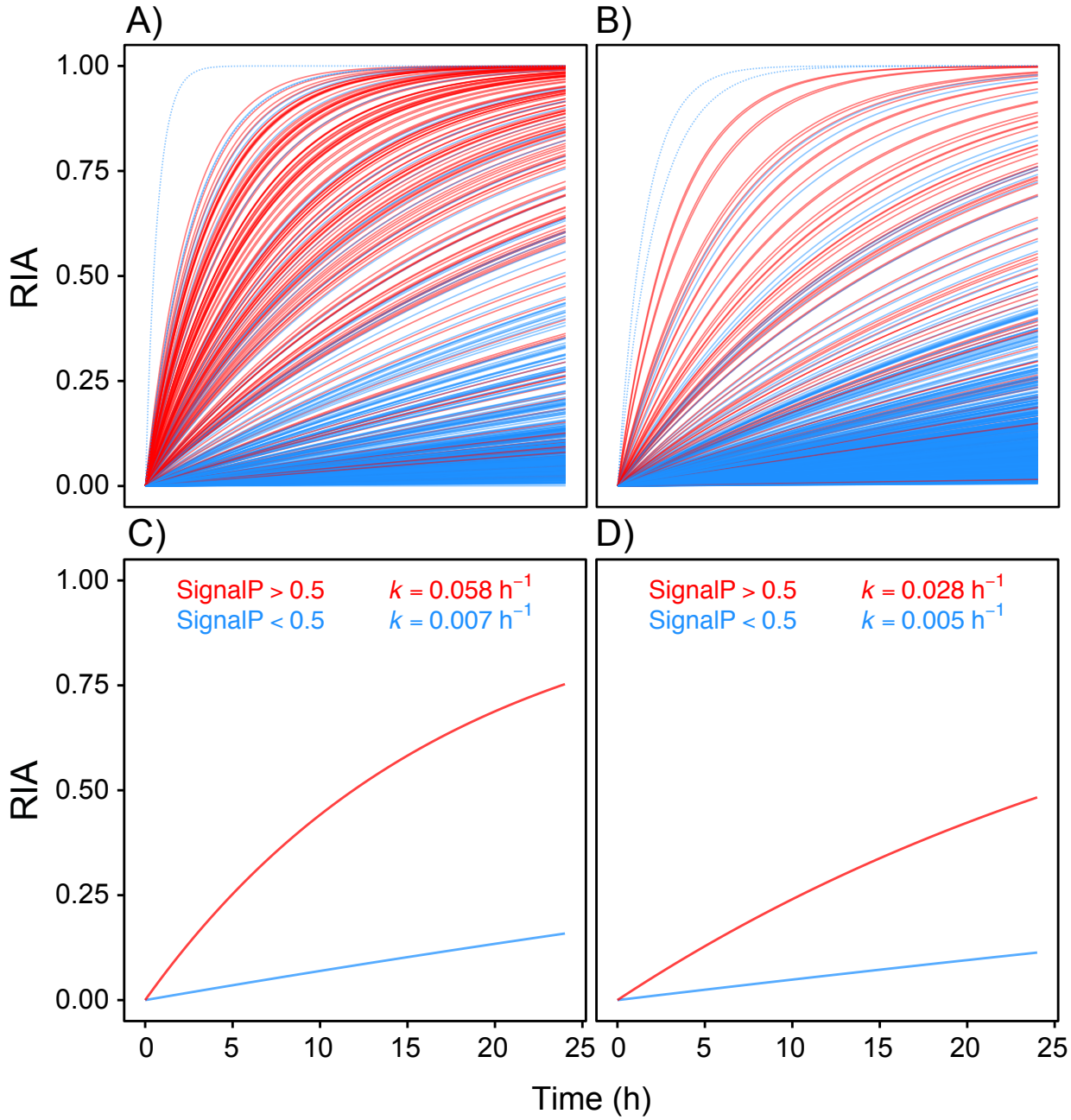


Figure 6

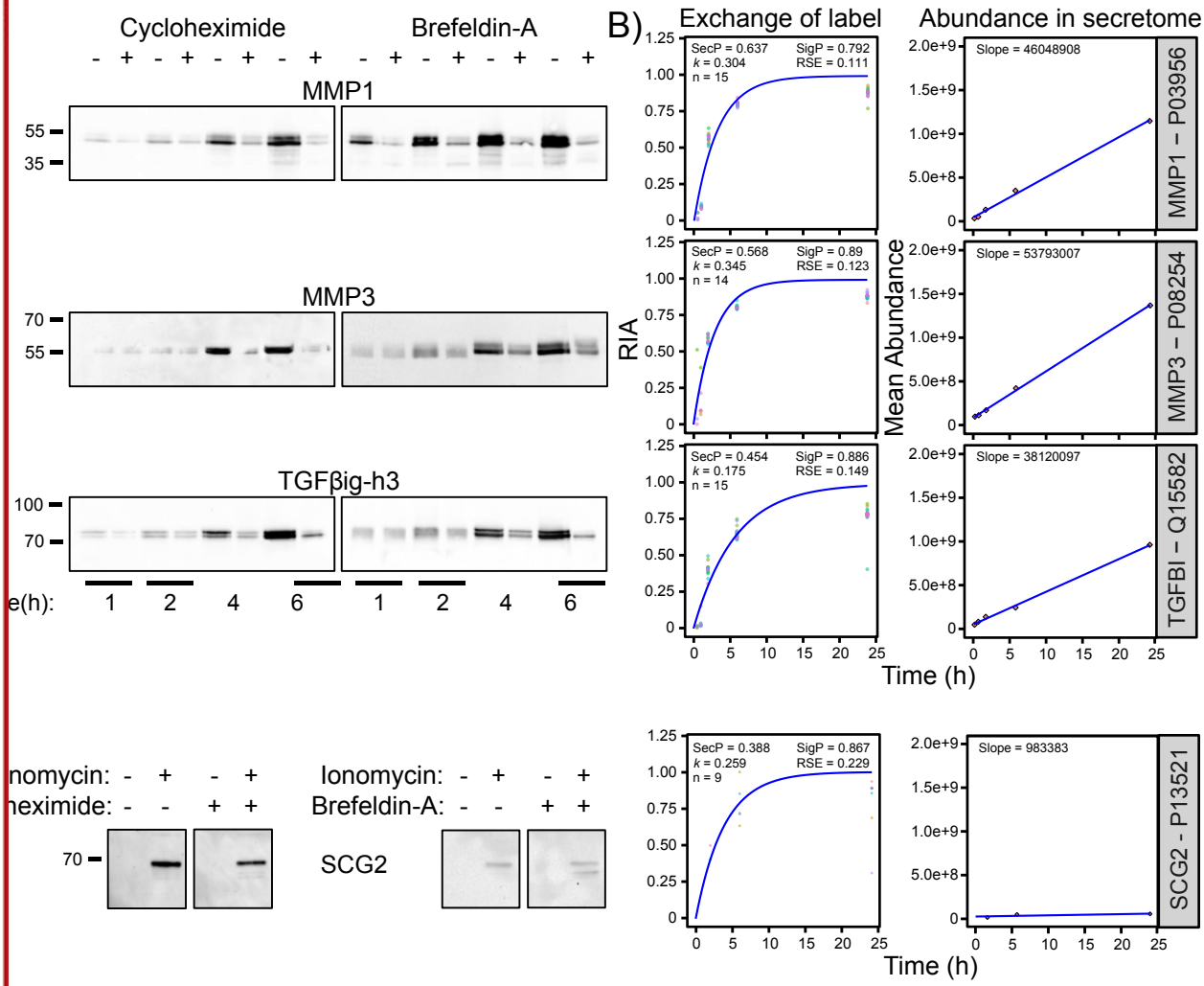

Cycloheximide
Brefeldin-A
- + - + - + - + - + - + - + - +
MMP1
55
35
MMP3
70
55
TGFβig-h3
100
70
e(h): 1 2 4 6 1 2 4 6
nomycin: - + - +
heximide: - - + +
70
Ionomycin: - + - +
Brefeldin-A: - - + +
SCG2
B)
Exchange of label
Abundance in secretome
SecP = 0.637
SigP = 0.792
k = 0.304
RSE = 0.111
n = 15
Slope = 46048908
MMP1 – P03956
SecP = 0.568
SigP = 0.89
k = 0.345
RSE = 0.123
n = 14
Slope = 53793007
MMP3 – P08254
SecP = 0.454
SigP = 0.886
k = 0.175
RSE = 0.149
n = 15
Slope = 38120097
TGFBI – Q15582
SecP = 0.388
SigP = 0.867
k = 0.259
RSE = 0.229
n = 9
Slope = 983383
SCG2 - P13521
RIA
Mean Abundance
Time (h)
Time (h)

Figure 7

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

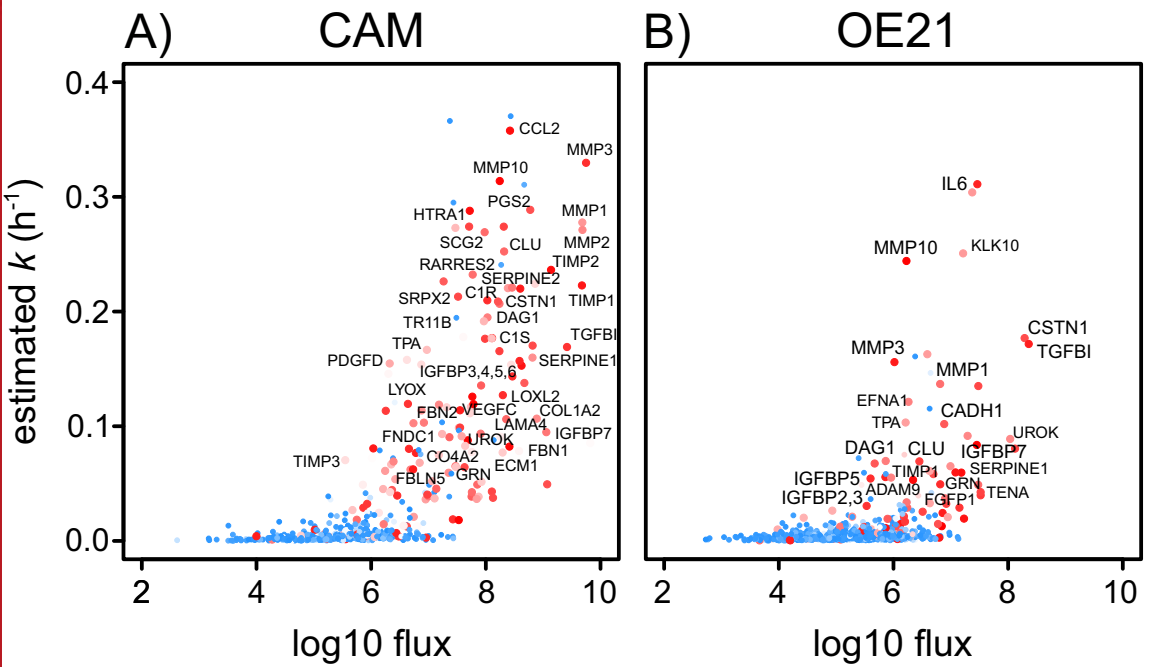


Figure 8

ASBMB
MOLECULAR & CELLULAR PROTEOMICS
MCP

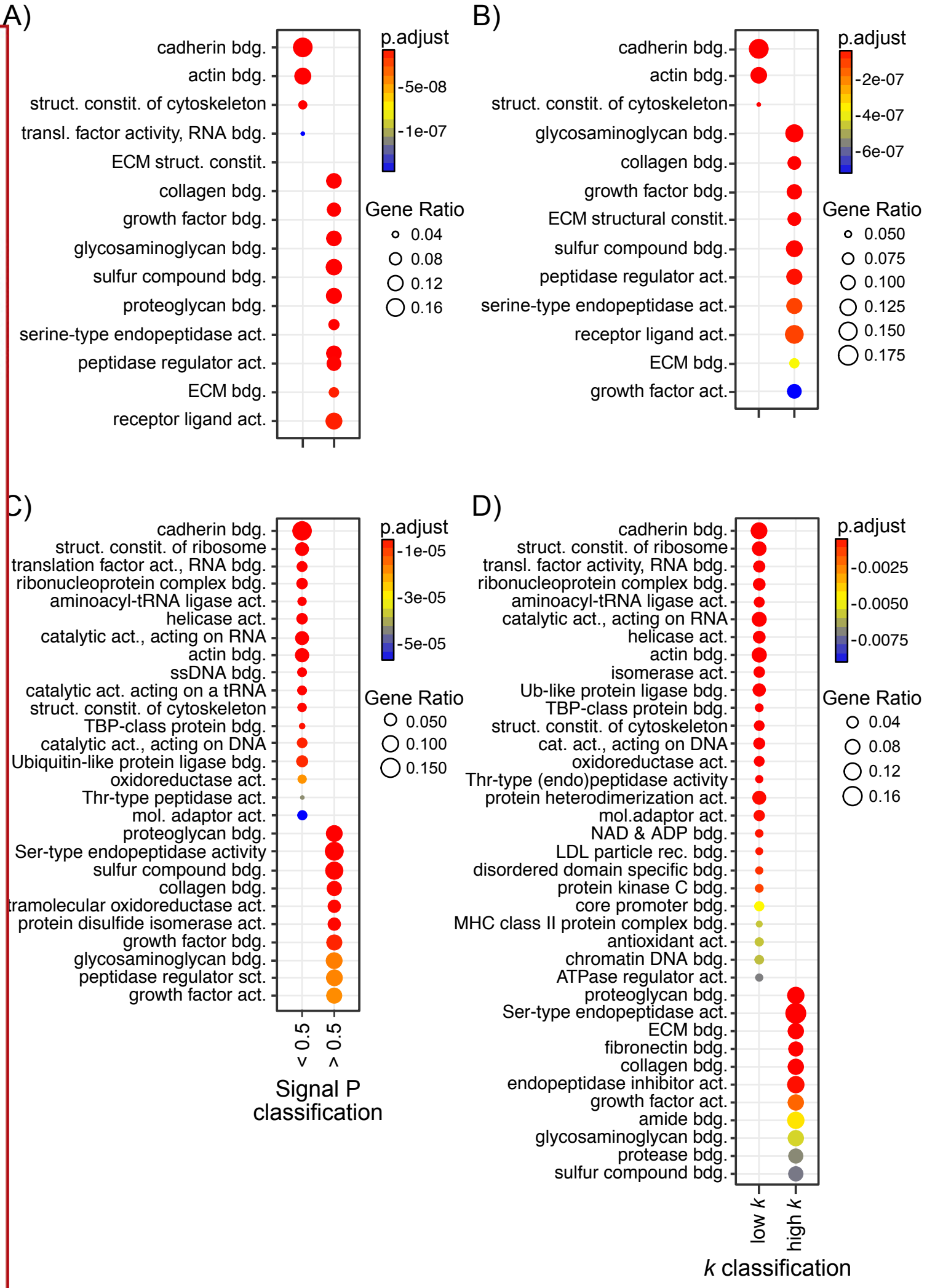


Figure 9