

Cell-selective labeling using amino acid precursors for proteomic studies of multicellular environments

Nicholas P Gauthier^{1,2}, Boumediene Soufi³, William E Walkowicz^{2,4}, Virginia A Pedicord⁵, Konstantinos J Mavrakis⁶, Boris Macek³, David Y Gin^{4,7}, Chris Sander¹ & Martin L Miller¹

We report a technique to selectively and continuously label the proteomes of individual cell types in coculture, named cell type-specific labeling using amino acid precursors (CTAP). Through transgenic expression of exogenous amino acid biosynthesis enzymes, vertebrate cells overcome their dependence on supplemented essential amino acids and can be selectively labeled through metabolic incorporation of amino acids produced from heavy isotope-labeled precursors. When testing CTAP in several human and mouse cell lines, we could differentially label the proteomes of distinct cell populations in coculture and determine the relative expression of proteins by quantitative mass spectrometry. In addition, using CTAP we identified the cell of origin of extracellular proteins secreted from cells in coculture. We believe that this method, which allows linking of proteins to their cell source, will be useful in studies of cell-cell communication and potentially for discovery of biomarkers.

The development and maintenance of multicellular environments is dependent on extensive cell-cell communication, and dysregulation of cellular interactions has a role in many diseases, including cancer^{1–3}. Although widely used for investigating intercellular signal transduction, antibody-based assays are relatively low throughput, vary in specificity and require preselection of protein readout. Although some of these limitations could be overcome using quantitative mass spectrometry-based proteomics^{4–6}, studies of cell-cell communication by mass spectrometry are hindered by an inability to distinguish between proteins from distinct cell types in multicell culture.

Recently several efforts have been made to differentiate the proteomes of individual cell populations in coculture. In one approach, each distinct cell type is labeled in isolation (for example, using heavy stable isotope-labeled L-lysine and/or L-arginine), and the fully labeled cells are subsequently mixed. Peptides identified with liquid chromatography-tandem mass spectrometry (LC-MS/MS) can then be assigned a source cell type on the basis of the status of the isotopic label. Two recent reports demonstrate the feasibility of such an approach for identifying early

ephrin signaling responses⁷ and determining proteins transferred between cell types⁸. Unfortunately, these labels become rapidly diluted as cells grow and divide in coculture, making this experimental setup primarily useful for investigating very early signaling events. In a different approach, differences in protein sequence between species are used to determine cell of origin in cross-species cocultures and xenografts^{9,10}. Although this approach can be used to distinguish between proteins from different cell types, the major drawbacks are that only a subset of peptides can be differentiated, and the findings from mixed-species models may not be physiologically relevant. Yet another approach relies on tRNA synthetases that specifically recognize and incorporate noncanonical amino acids into proteins^{11–13}. This method provides for both proteomic incorporation that is specific to transgenic cells as well as the ability to perform affinity enrichment on chemical moieties (for example, azides). However, structural differences between noncanonical and canonical amino acids might cause unpredictable functional alterations in mature proteins¹⁴. Given the caveats of each of these methods, there remains a need for a technique that enables continuous cell-specific labeling with canonical amino acids.

We report a method, CTAP, for cell-selective proteomic labeling that overcomes the limitations mentioned above. This method exploits the inability of vertebrate cells to synthesize certain amino acids required for growth and homeostasis. These ‘essential’ amino acids are produced in some plants, bacteria and lower eukaryotes, and must be supplemented in the medium of cultured vertebrate cells or obtained in the diet of animals¹⁵. We reasoned that transgenic expression of enzymes that synthesize essential amino acids would allow vertebrate cells to overcome auxotrophy by producing their own amino acids from supplemented precursors. These precursors can be isotopically labeled, allowing cell of origin of proteins to be determined by label status identified with MS/MS. We tested the validity and feasibility of CTAP using L-lysine, an essential amino acid commonly used in quantitative proteomic methods such as stable-isotope labeling by amino acids in cell culture (SILAC)^{5,16}. Using CTAP, we could continuously

¹Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. ²Louis V. Gerstner Jr. Graduate School of Biomedical Sciences, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. ³Proteome Center Tübingen, University of Tübingen, Tübingen, Germany. ⁴Molecular Pharmacology and Chemistry Program, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. ⁵Department of Immunology, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. ⁶Cancer Biology and Genetics Program, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. ⁷Deceased. Correspondence should be addressed to N.P.G. (ctap@cbio.mskcc.org), C.S. (ctap@cbio.mskcc.org) or M.L.M. (ctap@cbio.mskcc.org).

RECEIVED 10 JULY 2012; ACCEPTED 8 MAY 2013; PUBLISHED ONLINE 30 JUNE 2013; DOI:10.1038/NMETH.2529

Figure 1 | Overview of CTAP. (a) Schematic depicting the inability of vertebrate cells to grow on essential amino acids, including L-lysine, and that transgenic expression of L-lysine biosynthesis enzymes enables cells to produce and grow on their own supply of L-lysine. (b) Cells engineered in this way can be labeled selectively by supplementing the media with heavy isotope-labeled forms of the precursors. Expressing distinct L-lysine biosynthesis enzymes in different cell types enables continuous cell-selective proteome labeling using differentially labeled precursors when grown in medium lacking L-lysine. Quantitative mass spectrometry-based analysis of coculture samples can be used to determine relative abundance of proteins from each cell type. (c) CTAP can be used to investigate direct cell-cell contact or secreted factor mediated cell-cell communication, relevant for a range of biological phenomena.

and differentially label the proteome of cells in coculture, determine relative protein expression between the cell populations and identify the cell of origin of secreted factors.

RESULTS

Engineering mammalian cells to grow on L-lysine precursors

Several enzymes have been found in bacteria, fungi and plants that catalyze reactions leading to the production of L-lysine from precursor compounds. We hypothesized that by engineering vertebrate cells to produce their own supply of L-lysine from labeled precursors, we could achieve differential proteomic labeling of specific cell types in coculture (Fig. 1). We began by identifying precursor-enzyme pairs in which the precursor was readily available and the enzyme had no described orthologs in vertebrate genomes (Supplementary Fig. 1). In a different context, one of the candidate precursor-enzyme pairs had been used to rescue L-lysine auxotrophy when creating a positive selection system for vector incorporation^{17,18}. To investigate the candidate precursors and eliminate those that autonomously rescue L-lysine

auxotrophy, we examined growth rates in SILAC media supplemented with L-lysine, various precursors or in L-lysine-free conditions. With the exception of N^{α} -acetyl-L-lysine, the tested precursors alone had little or no effect on growth in wild-type cells (Supplementary Fig. 2).

We next investigated whether transgenic expression of enzymes involved in biosynthesis of L-lysine would allow cells to acquire the ability to grow on precursors. We stably expressed genes encoding the enzymes lysine racemase (lyr) from *Proteus mirabilis* (optimized for expression and intracellular localization; Supplementary Note 1) and diaminopimelate decarboxylase (DDC) from *Arabidopsis thaliana* in several cell lines (Supplementary Table 1). DDC-expressing mouse 3T3 and HEK293T cells, along with lyr-expressing human MDA-MB-231 cells, exhibited growth rates in media supplemented with the precursors 2,6-diaminopimelic acid (DAP) and D-lysine, respectively, comparable to those in medium containing L-lysine (Fig. 2a,b and Supplementary Fig. 3). We observed similar results for several additional human and mouse cell lines, although two

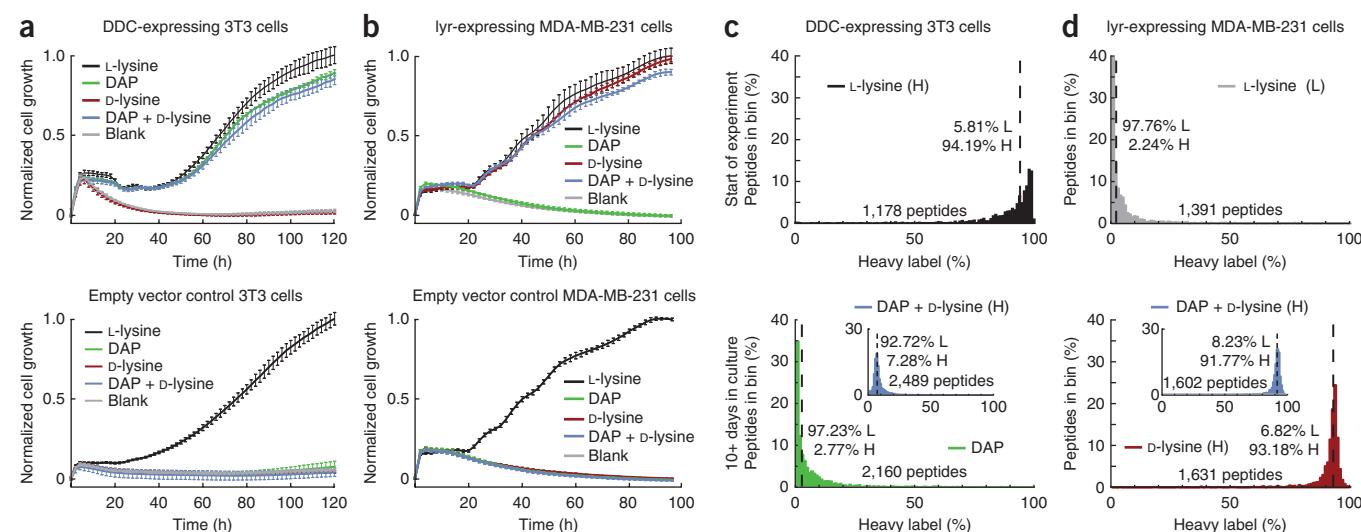


Figure 2 | Vertebrate cell lines expressing L-lysine biosynthesis enzymes grow on and incorporate L-lysine produced from their precursors. (a,b) Cell growth, assessed with impedance (a correlate of the number of cells) using the xCELLigence system (normalized to maximum growth) for mouse fibroblast 3T3 cells that stably express DDC (a) and human breast carcinoma MDA-MB-231 cells that stably express lyr plated in L-lysine-free medium supplemented with 10 mM DAP, 4 mM D-lysine, both precursors or 0.798 mM L-lysine. Data for control cells (with empty vector) are shown (bottom). Error bars, s.d. of three biological replicates. (c,d) Label status of lysine-containing peptides assessed by quantitative LC-MS/MS of cell lysates collected at the start of the experiment from monocultured DDC-expressing 3T3 cells labeled with heavy L-lysine (H; c) and lyr-expressing MDA-MB-231 cells labeled with light L-lysine (L; d) (top). Percentage incorporation of heavy label was determined using H/L ratios using MaxQuant software. Same analyses for cells that were further cultured ten or more days (10+; two passages) in L-lysine-free medium containing the indicated precursors (bottom). Dashed black lines indicate median peptide (percentages indicated). Each histogram depicts one sample processed by LC-MS/MS.

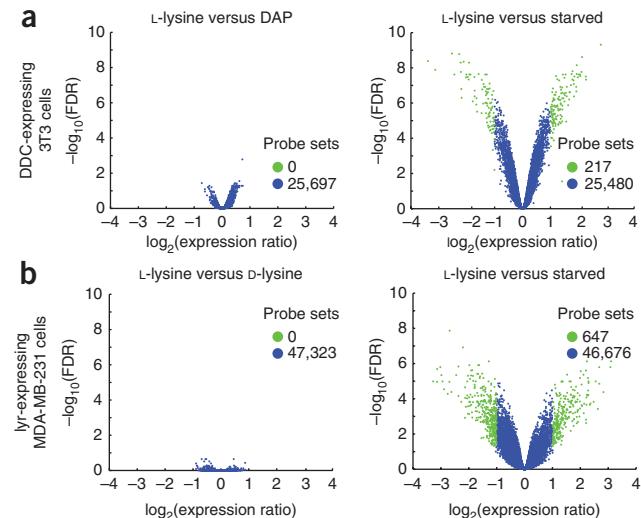
Figure 3 | Limited gene expression changes observed when growing cells in precursor versus L-lysine. **(a)** Gene expression profiled using the Illumina microarray platform on mRNA collected 72 h after DDC-expressing 3T3 cells were plated in SILAC medium supplemented with DAP, L-lysine or neither (starved). Expression differences of DAP versus L-lysine (left) and starved versus L-lysine (right) are plotted as a function of significance (moderated *t*-statistics adjusted for multiple testing by the Benjamini and Hochberg method). Highlighted genes (green) are more than twofold differentially regulated at the level of FDR < 0.05. **(b)** As in **a**, except MDA-MB-231 cells expressing lyr were plated on L-lysine, D-lysine or in starved conditions. All experiments were performed in triplicate.

of the DDC-expressing lines never reached growth rates comparable to those in the presence of L-lysine (**Supplementary Table 2**). The enzyme-precursor pairs were specific, as we observed no growth with mismatched enzyme-precursor pairs or in empty-vector controls (**Fig. 2a,b**). Furthermore, growth in standard conditions with L-lysine and the addition of either DAP or D-lysine resulted in little or no growth perturbation (**Supplementary Fig. 4**). These monoculture results show that transgenic enzyme expression with supplementation of specific precursors is responsible for the growth rescue observed in L-lysine-free conditions.

Proteomic incorporation of precursor-based L-lysine

To investigate whether L-lysine is directly produced by enzymatic turnover of the supplemented precursors, we applied the SILAC principle of exchanging the label of amino acids from one form to another (for example, light L-lysine to heavy L-lysine)⁵. At the beginning of the experiments, we labeled DDC-expressing 3T3 cells with heavy [¹³C₆,¹⁵N₂]L-lysine (H) and labeled lyr-expressing MDA-MB-231 cells with light L-lysine (L). Then we grew these cells in monoculture for ten or more days (two or three passages) in L-lysine-free media that contained light isotope-labeled DAP (L), heavy-labeled [²H₈]D-lysine (H) or both precursors. We digested protein from cell lysates with trypsin and/or LysC, submitted the products to high-resolution LC-MS/MS and determined the H/L ratio for each peptide using the MaxQuant software package¹⁹. We used these H/L ratios to determine the labeling incorporation levels, which we found to be similar to orthogonal enrichment calculations (**Supplementary Table 3**) and robust to a range of parameter changes, such as peptide score and length (**Supplementary Table 4**).

In the presence of light isotope-labeled DAP alone, peptides identified in DDC-expressing 3T3 cells switched from being predominantly labeled with heavy L-lysine (94% H, median peptide) to being labeled with light L-lysine (97% L) (**Fig. 2c**; all peptide and protein lists with associated H/L ratios are provided in **Supplementary Data**). Similarly, the peptides identified in lyr-expressing MDA-MB-231 cells changed from 98% L to 93% H in the presence of heavy isotope-labeled D-lysine (**Fig. 2d**). This level of labeling can be reached after one passage (approximately four doublings, **Supplementary Fig. 5**) and can be considered near complete as it is similar to that in the initial samples and what is typically reported in SILAC experiments^{20,21}. To test the amount of unspecific labeling (cross-contamination), we grew cultures in the presence of both precursors. Supplementing the DDC precursor DAP (L) had little effect on the label switch in lyr-expressing MDA-MB-231 cells, but the presence of D-lysine (H) marginally increased the amount of heavy label in DDC-expressing



3T3 cells (**Fig. 2c,d**). This difference in the 3T3 cells was possibly due to contamination of heavy L-lysine in heavy D-lysine (currently only ≥95% enantiomeric purity in reagent from C/D/N Isotopes) and is expected to be lower with greater purity. Taken together, these data indicate that lyr- and DDC-expressing cells can specifically incorporate L-lysine synthesized directly from these enzymes' respective precursors.

Limited perturbation to cells growing on precursors

We next investigated whether cells behave similarly when grown on precursors compared to L-lysine. We cultured cells for 3 d in media containing L-lysine, precursor or neither (starved, positive control for perturbed state) and profiled mRNA expression using microarrays (**Fig. 3** and **Supplementary Fig. 6**). Relative to the basal L-lysine condition, expression of no genes changed substantially when we grew enzyme-expressing cells in the presence of precursor, but expression of hundreds of genes changed in starvation conditions (false discovery rate (FDR) < 0.05 and expression ratio > 2; **Fig. 3**). Furthermore, we performed several assays to probe the effects of precursor-based growth, including gene-set enrichment analysis (**Supplementary Tables 5** and **6**), measurement of amino acid starvation factors (**Supplementary Fig. 7**), determination of protein abundance by LC-MS/MS (**Supplementary Fig. 8**) as well as growth and molecular response to drug perturbation (**Supplementary Figs. 9** and **10**). Although we observed minor differences, overall these data demonstrate that growing cells with the precursors specific to the expressed enzyme had little to no effect compared to growth on L-lysine.

Continuous and differential proteome labeling in coculture

To assess the specificity of labeling in coculture with each cell population using a distinct enzyme-precursor pair, we took advantage of species-specific sequence differences to compare label status between the enzyme-expressing mouse 3T3 and human MDA-MB-231 cell lines. Initially, we grew the 3T3 cells in heavy L-lysine (H) and the MDA-MB-231 cells in light L-lysine (L). At the start of the experiment, we collected a sample from each cell line and combined the cell pellets 1:1 to verify the ability to differentiate label status based on species-specific peptide classification. As expected, we confirmed labels of mouse-specific and

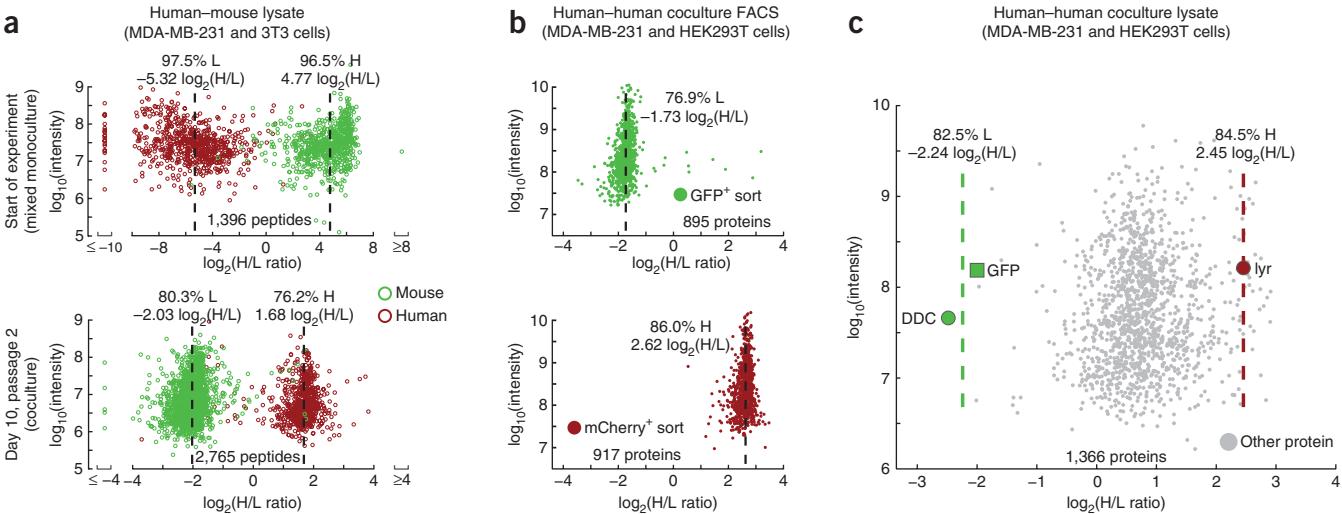


Figure 4 | Using two distinct enzyme-precursor pairs, cocultured cells exhibit precursor-based differential proteome labeling. **(a)** LC-MS/MS analysis of a mixture of DDC-expressing 3T3 cells (mouse) labeled with heavy isotope-labeled L-lysine (H) and lyr-expressing MDA-MB-231 cells (human) labeled with light isotope-labeled L-lysine (L) (mixed monoculture, top). Similarly labeled cells were then cocultured and analyzed after 10 d (two passages) on DAP (L) and D-lysine (H) (bottom). Median indicated by dashed black lines. **(b)** LC-MS/MS analysis of sorted GFP⁺ (top) and mCherry⁺ (bottom) samples from cocultured DDC-expressing GFP⁺ HEK293T and lyr-expressing mCherry⁺ MDA-MB-231 cells. Before analysis the coculture was grown for 5 d (one passage, approximately four cellular doublings) in medium containing DAP (L) and D-lysine (H). **(c)** Proteins derived from unsorted coculture of cells as in **b**. Mean of H/L ratios of the transgenes specific to each cell population (DDC and GFP in HEK293T cells and lyr in MDA-MB-231 cells) are indicated with green and red lines. Note that although mCherry was detected, it is not included as a distinct protein as it is fused to lyr. Each plot depicts one sample processed by LC-MS/MS.

human-specific peptides to be primarily heavy and light, respectively (Fig. 4a, Supplementary Fig. 11 and Supplementary Data).

With the expectation that each cell type would exchange label, then we combined the prelabeled cells in coculture in medium containing 10 mM light DAP (L) and 1 mM heavy D-lysine (H). After 10 d (two passages), the two cell types switched labels (Fig. 4a and Supplementary Fig. 12). As expected, the mouse 3T3 peptides became predominantly labeled with light lysine (80% or $-2.0 \log_2 H/L$) and the human MDA-MB-231 peptides became predominantly labeled with heavy lysine (76% or $1.7 \log_2 H/L$). This labeling efficiency is lower than in the mixed monoculture control and could either be due to physiological sharing of L-lysine between cell populations or limitations of the CTAP method. Although we optimized CTAP to remove shared L-lysine from the medium with daily changes of the medium and modified the lyr sequence to decrease enzyme secretion (for example, removal of signal peptide and addition of intracellular targeting sequence, Supplementary Fig. 13), we expect that future development will increase enrichment of cell type-specific label even further. Despite the fact that the absolute H/L ratios of the species-specific peptides differed between the mixed monoculture control and the CTAP-labeled coculture, the human and mouse-specific peptides were clearly separable in both experiments. These distinct H/L ratios in species-specific sequences therefore demonstrate the ability to differentially label the proteome across cell types in coculture.

We next investigated whether the CTAP method could be used to differentiate the proteome of a same-species coculture system. We plated DDC-expressing GFP⁺ HEK293T cells together with lyr-expressing monomeric (m)Cherry⁺ MDA-MB-231 cells. After 5 d (one passage, approximately four cell doublings) of growth in light isotope-labeled DAP (L) and heavy isotope-labeled D-lysine (H), we sorted a coculture sample for

mCherry⁺ and GFP⁺ cells by fluorescence-activated cell sorting (FACS; Supplementary Fig. 14) and separately subjected each of the sorted populations to LC-MS/MS. Analysis of protein from the GFP⁺ and mCherry⁺ cells of this human-human coculture revealed labeling efficiency similar to that seen in the human-mouse coculture, with each cell population exhibiting distinct H/L ratios (Fig. 4b and Supplementary Data).

We collected another sample directly from the unsorted human-human coculture and identified 1,366 proteins with LC-MS/MS (Supplementary Data). Focusing on the transgenic proteins exclusive to each cell population (GFP and DDC for the HEK293T as well as lyr in MDA-MB-231 cells), we observed the expected H/L ratios corresponding to those determined by FACS (Fig. 4c). This concordance confirms differential labeling in human-human coculture lysates. When analyzing all identified proteins in this unsorted sample, the H/L ratios exhibited a near-normal distribution with the transgenes lying in the tails of the distribution. Although these tails contain relatively few members, they likely represent cell type-specific proteins (Fig. 4c and Supplementary Fig. 15). This result is consistent with a recent report that found most proteins are ubiquitously expressed across different cell types but with different relative abundance²². In summary, these results demonstrate the capability of CTAP to label the proteome in a cell-specific manner and show that label status (H/L ratio) is directly related to the relative protein abundance level between the two cell types.

Linking secreted proteins to their cell of origin

To test the potential of using the CTAP method to discriminate the cell of origin of secreted factors, we collected supernatant from the same human and mouse coculture setup as in the previous experiments. We grew the cells for 16 h in serum-free medium to avoid overloading the sample with serum proteins,

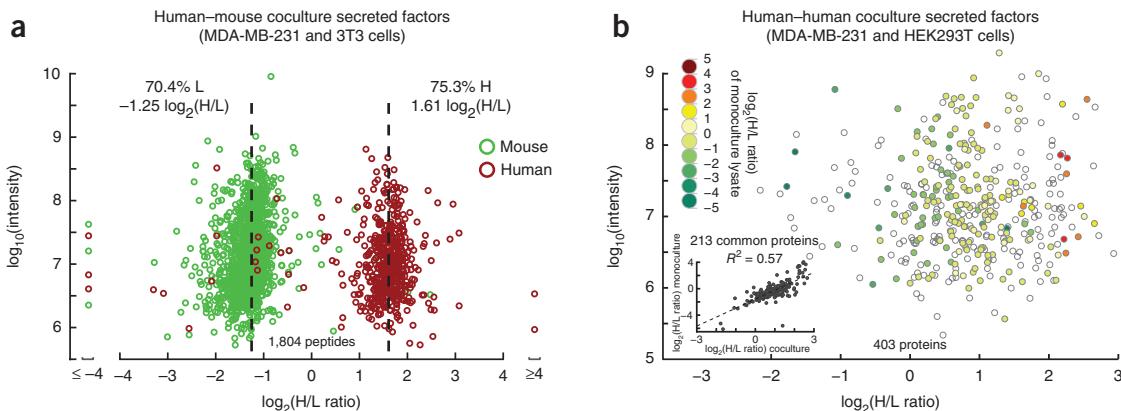


Figure 5 | Application of CTAP for determining cell of origin for secreted factors. **(a)** LC-MS/MS analysis of proteins in the supernatant of DDC-expressing 3T3 cells (mouse) and lyr-expressing MDA-MB-231 cells (human) cocultured in DAP (L) and α -lysine (H). Before sample collection, cells were grown for 16 h in serum-free medium. Only peptides that are unique to mouse and human are shown. **(b)** Similar to **a**, except coculture consisted of two human cell lines: HEK293T cells expressing DDC and MDA-MB-231 cells expressing lyr. Colors depict relative protein abundance as determined by SILAC quantitation of mixed, separately labeled monoculture lysates. Inset, correlation between the H/L ratios for the same monoculture and coculture samples. Uncolored points represent proteins that were not identified in the monoculture sample. Each plot depicts one sample processed by LC-MS/MS.

and then collected the supernatant. We concentrated the secreted proteins by ultracentrifugation, precipitated them with methanol and chloroform, and analyzed them by LC-MS/MS. Nearly all species-specific peptides could be distinguished by label alone (**Fig. 5a**, **Supplementary Fig. 16** and **Supplementary Data**). These results demonstrate the capability of the method to determine the cell of origin for secreted proteins in coculture.

Applying a similar approach to analyze secreted factors in a same-species coculture, we collected supernatant from the same cocultured DDC-expressing HEK293T and lyr-expressing MDA-MB-231 cells as used previously and subjected it to LC-MS/MS. The H/L ratios of 403 identified proteins spanned a similar range as those detected in cells (**Figs. 4c**, **5b** and **Supplementary Data**). Having shown that the H/L ratios are distinct for species-specific proteins in the human and mouse coculture secretome, the proteins in the tails of this human-human distribution likely are cell type-specific. Reasoning that a protein secreted primarily by one cell type would also be relatively more abundant in cells, we investigated whether CTAP-labeled extracellular protein ratios correlated with SILAC-labeled protein ratios obtained from mixed monoculture lysates. For the subset of proteins that were common to both samples, we observed good agreement between the H/L ratios from the combined intracellular monocultures and secreted coculture samples ($R^2 = 0.57$, **Fig. 5b** and **Supplementary Fig. 17**). The positive correlation between these conditions indicates that the secreted proteins with the lowest and highest H/L ratios are likely cell type-specific as they are most abundant in the HEK293T and MDA-MB-231 cells, respectively (**Supplementary Data**). Almost half of the putative MDA-MB-231 cell secreted proteins were not identified intracellularly (47%), highlighting the need for secretome profiling. Taken together with the species-verified secretome analysis, these results establish that the CTAP method can be applied to determine the cell of origin of secreted factors in coculture.

DISCUSSION

To our knowledge, CTAP is the only method in which the proteome of specific cell populations can be labeled continuously and differentially by canonical amino acids in a complex mixture of cells.

Although the results of this initial study demonstrate the feasibility and functionality of CTAP, there are several avenues for development of this method.

We have preliminary data showing that CTAP can be applied to label a specific cell type of interest in a mixed cell culture system using only one enzyme-precursor pair (**Supplementary Fig. 18**; note that titrating down the amount of L-lysine in the medium was required). A one-enzyme approach may facilitate the linking of proteins to a distinct cell population in experimental systems where a two-enzyme approach is either not desirable or not feasible. In addition to using DDC and lyr, we have also tested and found specific but suboptimal growth rescue with the enzyme CBZcleaver and substrate Z-lysine, supporting the general principle of this method (**Supplementary Fig. 19** and **Supplementary Note 2**). Future method development will explore the use of single-enzyme labeling in complex cultures as well as focus on the identification of additional enzyme-precursor pairs.

The coculture labeling efficiency was lower than expected from mixed monoculture controls (approximately 80% versus 97%, **Fig. 4**). Incomplete labeling in coculture can arise from both technical problems and true biological interactions between the two cell populations. During method development, we identified extracellular lyr activity as a potential source of label dilution. Even though we implemented several experimental optimizations (**Supplementary Fig. 13**), extracellular lyr activity likely still has a role in coculture label contamination, and our future efforts will focus on complete suppression of this activity. Other optimization steps will involve improving enzyme efficacy, increasing precursor uptake^{17,18} (**Supplementary Discussion**) and increasing purity of heavy α -lysine (currently only $\geq 95\%$ enantiomeric pure as supplied by C/D/N Isotopes). Despite technical improvements, biological exchange of L-lysine between cells in coculture may prevent cell type specific-labeling from reaching isotopic enrichment as complete as that in SILAC. Nevertheless, the distinct H/L ratios we observed in each cell population readily enable clear identification of cell type-specific factors and relative quantitation of protein expression between the populations.

Several criteria must be met for cells to qualify for the use of CTAP to study cell-cell communication. Similar to SILAC, cells

must be L-lysine auxotrophic and able to grow in dialyzed serum or serum-free media. Currently, the method relies on precursors of L-lysine only, hampering the use of L-arginine containing peptides for protein quantitation of tryptic digests. Additionally, CTAP requires stable expression of exogenous enzymes and growth on amino acid precursors, and investigators should verify that the context-specific phenotype of their cells is unperturbed when using this method. Although the H/L ratios of the transgenic proteins themselves can be used to determine the level of labeling in each cell population, it is advisable that the incorporation efficiency be verified by LC-MS/MS analysis of samples subjected to FACS. These levels can be used as cutoffs for assigning the cell of origin to individual proteins. Unequal cell number may influence labeling efficiency in coculture due to sharing of L-lysine from residual extracellular Lys activity or physiological amino acid exchange. This potential issue can be addressed by using different seeding densities of each cell population and daily changes of the culture medium. Despite incomplete labeling, normalizing the H/L ratios (for example, using Z scores) can enable relative comparison of protein changes between conditions in a cell type-specific manner.

Upon overcoming some of the remaining technical challenges, such as competition with endogenous L-lysine, precursor delivery and enzyme expression, another possible application for CTAP may be identification of disease biomarkers *in vivo*. Current approaches for biomarker identification are limited by their inability to classify whether a potential marker originates from the diseased tissue itself or from the normal tissue. Using the described technique we can potentially circumvent these limitations, as proteins from specific cell types of interest can, in principle, be labeled continuously *in vivo*. Any labeled protein identified in the serum or proximal fluids will have originated from the cell type of interest. As CTAP allows for unbiased and high-throughput LC-MS/MS to differentiate peptides derived from distinct cells in complex cellular environments, we anticipate that CTAP will be an important tool for gaining insight into intercellular signaling in a range of fundamental biological processes.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Accession codes. GenBank: [KC962560](#), [KC962561](#), [KC962562](#) and [KC962563](#) (oligonucleotide sequences for inserts). Gene Expression Omnibus: [GSE43894](#) and [GSE43895](#) (Illumina microarray experiments).

Note: Supplementary information is available in the [online version of the paper](#).

ACKNOWLEDGMENTS

We acknowledge E. Larsson, Y. Gruber, D.S. Marks, A. Arvey, J. Joyce and A. Koff for helpful discussions, H. Erdjument-Bromage for pilot MS/MS investigation, A.N. Miller, J. Cross and X. Jing for technical help, and E. Larsson, J. Gauthier, J. Joyce and A.M. Miller for helpful comments on the manuscript. This work was funded in part by US National Cancer Institute grant U54 CA148967.

AUTHOR CONTRIBUTIONS

N.P.G. and M.L.M. designed and performed experiments and analyzed data. W.E.W. generated reagents. B.S., K.J.M. and V.A.P. contributed to experiments. N.P.G. and M.L.M. wrote the manuscript. B.S., W.E.W., B.M., K.J.M., V.A.P., D.Y.G. and C.S.

contributed to discussions and editing of the manuscript. N.P.G. conceived the hypothesis. N.P.G., C.S. and M.L.M. developed the concept and managed the project.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the [online version of the paper](#).

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Joyce, J.A. & Pollard, J.W. Microenvironmental regulation of metastasis. *Nat. Rev. Cancer* **9**, 239–252 (2009).
- Mcmillan, D.W. *et al.* Tumor cell-specific bioluminescence platform to identify stroma-induced changes to anticancer drug activity. *Nat. Med.* **16**, 483–489 (2010).
- Olumi, A.F. *et al.* Carcinoma-associated fibroblasts direct tumor progression of initiated human prostatic epithelium. *Cancer Res.* **59**, 5002–5011 (1999).
- Gygi, S.P. *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999 (1999).
- Ong, S.-E. *et al.* 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 (2002).
- Ross, P.L. *et al.* Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell Proteomics* **3**, 1154–1169 (2004).
- Jørgensen, C. *et al.* Cell-specific information processing in segregating populations of Eph receptor ephrin-expressing cells. *Science* **326**, 1502–1509 (2009).
- Rechavi, O. *et al.* Trans-SILAC: sorting out the non-cell-autonomous proteome. *Nat. Methods* **7**, 923–927 (2010).
- Naba, A. *et al.* The matrisome: in silico definition and *in vivo* characterization by proteomics of normal and tumor extracellular matrices. *Mol. Cell Proteomics* **11**, M111.014647 (2012).
- van den Bemd, G.-J.C.M. *et al.* Mass spectrometric identification of human prostate cancer-derived proteins in serum of xenograft-bearing mice. *Mol. Cell Proteomics* **5**, 1830–1839 (2006).
- Ngo, J.T. *et al.* Cell-selective metabolic labeling of proteins. *Nat. Chem. Biol.* **5**, 715–717 (2009).
- Truong, F., Yoo, T.H., Lampo, T.J. & Tirrell, D.A. Two-strain, cell-selective protein labeling in mixed bacterial cultures. *J. Am. Chem. Soc.* **134**, 8551–8556 (2012).
- Ngo, J.T., Schuman, E.M. & Tirrell, D.A. Mutant methionyl-tRNA synthetase from bacteria enables site-selective N-terminal labeling of proteins expressed in mammalian cells. *Proc. Natl. Acad. Sci. USA* **110**, 4992–4997 (2013).
- Liu, C.C. & Schultz, P.G. Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **79**, 413–444 (2010).
- Xu, H., Andi, B., Qian, J., West, A.H. & Cook, P.F. The alpha-amino adipate pathway for lysine biosynthesis in fungi. *Cell Biochem. Biophys.* **46**, 43–64 (2006).
- Ong, S.-E. & Mann, M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat. Protoc.* **1**, 2650–2660 (2006).
- Saqib, K.M., Hay, S.M. & Rees, W.D. The expression of *Escherichia coli* diaminopimelate decarboxylase in mouse 3T3 cells. *Biochim. Biophys. Acta* **1219**, 398–404 (1994).
- Jouanneau, J., Stragier, P., Bouvier, J., Patte, J.C. & Yaniv, M. Expression in mammalian cells of the diaminopimelic acid decarboxylase of *Escherichia coli* permits cell growth in lysine-free medium. *Eur. J. Biochem.* **146**, 173–178 (1985).
- Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
- Sury, M.D., Chen, J.-X. & Selbach, M. The SILAC fly allows for accurate protein quantification *in vivo*. *Mol. Cell Proteomics* **9**, 2173–2183 (2010).
- Spellman, D.S., Deinhardt, K., Darie, C.C., Chao, M.V. & Neubert, T.A. Stable isotopic labeling by amino acids in cultured primary neurons: application to brain-derived neurotrophic factor-dependent phosphotyrosine-associated signaling. *Mol. Cell Proteomics* **7**, 1067–1076 (2008).
- Geiger, T., Wehner, A., Schaab, C., Cox, J. & Mann, M. Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Mol. Cell Proteomics* **11**, M111.014050 (2012).

ONLINE METHODS

Oligonucleotide acquisition. The L-lysine-producing enzymes used in this study were DDC, lyr and CBZcleaver. Sequence encoding DDC was directly amplified by PCR from *Arabidopsis thaliana* cDNA (TAIR identifier AT3G14390; primer sequences are available in **Supplementary Tables 7** and **8**). The lyr and CBZcleaver constructs were synthesized by GeneArt, encoding the amino acid sequences specified previously^{23,24}, and nucleotide sequences were optimized for expression in mouse. All sequences are available in **Supplementary Note 1**. Sequences were verified for all plasmids by the Sanger method of sequencing.

Vector cloning, viral production and cell-line creation. Two murine stem cell virus (MSCV)-based retroviral vector backbones, one expressing GFP (pMIG) and the other mCherry (pMIC), were used to infect mouse cells. To create the insert for pMIG, the PCR product of DDC sequence was cloned into the EcoRI site of the vector. CBZcleaver sequence was directly subcloned from the GeneArt-supplied vector pMA-RQ into pMIC using EcoRI and XhoI restriction sites. Viral supernatants for pMIG and pMIC were produced by transfecting Phoenix cells with each plasmid, and the supernatant was used to infect 3T3 cells 48 h later as previously described^{25,26}.

The lentiviral backbone pLM was used to infect human cells in this study. Overlapping PCR was performed to generate sequences encoding GFP-DDC and mCherry-(wild type)lyr constructs that were linked by a sequence encoding P2A peptide preceded by a Gly-Ser-Gly linker²⁷. Overlapping and insertional PCR was used to generate a truncated form of lyr that was fused to mCherry and a mitochondrial targeting sequence (**Supplementary Table 7**). The pLM-enzyme virus was packaged by calcium phosphate transfection of the HEK293T packaging cell line using 10 µg of transfer vector, 6.5 µg of CMVδR8.74 and 3.5 µg of the VSV.G plasmid. MDA-MB-231 and HEK293T cells were then infected with lentiviral supernatant produced from the pLM construct 48 h after transfection of the packaging line.

Cellular growth assays. Cell lines were grown in Dulbecco's modified Eagle's medium (DMEM) without L-lysine and L-arginine (SILAC DMEM, Thermo Fisher Scientific) supplemented with 10% dialyzed FBS (Sigma, F0392), antibiotics and L-glutamine. For monoculture growth assays, 1 mM L-arginine was added to the medium, and 4,000 or 5,000 cells were seeded in 200 µl medium per well in 96-well plates with various concentrations of L-lysine, 2,6-diaminopimelic acid (DAP, Sigma, 33240), D-lysine HCl (Sigma, L5876), N^α-Cbz-L-lysine (Z-lysine, BaChem, C-2200) or N^α-acetyl-L-lysine (N2A, Sigma, A2010). Although the substrate of DDC is only the meso form of DAP, in this work the DAP used contains D, L and meso forms. Cell viability was measured using either the metabolic activity-based Resazurin (Sigma) reagent or the impedance-based xCELLigence system (Roche). For Resazurin experiments, 25 µl of the Resazurin reagent was added to each well, and cellular growth was estimated after 2–3 h of incubation at 37 °C as described by the manufacturer. For xCELLigence experiments, cells were plated in either 16-well or 96-well E-plates, allowed to settle for 30 min at room temperature, and then placed in the xCELLigence Real Time Cell Analyzer Dual-Plate (RTCA DP) or Real Time Cell Analyzer Multiple-Plate (RTCA MP) ana-

lyzer, where impedance was measured every 15 min for 96–120 h. At least three replicates were performed for each condition.

Percentage of mCherry⁺ and GFP⁺ cells in coculture was measured by either flow cytometry (BD LSR II) or Tali image-based cytometry (Invitrogen). For flow cytometry assays, 25,000 cells from each cell line were seeded together in 6-well plates in 3–4 ml of medium supplemented with different concentrations of L-lysine and/or L-lysine precursors. After 72 h, cells were trypsinized, washed and resuspended in 200 µl PBS containing 2% dialyzed FBS and 0.1% NaN₃. For estimating total cell numbers using the ViaCount assay (Millipore), 20 µl of the resuspended sample was used as described by the manufacturer. The remaining 180 µl was mixed with an equal volume of 2% paraformaldehyde and the percentage of GFP⁺ and mCherry⁺ cells in each sample was analyzed by flow cytometry. At least two replicates were performed for each condition. For Tali assays, cells were trypsinized and resuspended in medium; 25 µl of coculture cell suspension was used to determine the percentage of GFP⁺ and RFP⁺ cells in biological triplicates.

Stable-isotope labeling and cell passaging. For exchange-of-label experiments (all monocultures, all human and mouse cocultures and experiments in **Supplementary Fig. 19**), cells were first metabolically labeled by growth for at least 10 cellular doublings (ten or more days) in L-arginine and 10% dialyzed FBS-containing SILAC DMEM supplemented with 798 µM light isotope-labeled L-lysine (L), medium isotope-labeled [²H]₄L-lysine (M, +4 daltons) or heavy isotope-labeled [¹³C₆, ¹⁵N₂]L-lysine (H, +8 daltons; Cambridge Isotopes). Cells were then seeded in monoculture or coculture with 10 mM light DAP (L, Sigma), 1 mM or 4 mM heavy isotope-labeled [²H]₈D-lysine (H, +8 daltons; C/D/N Isotopes; 3,3,4,4,5,5,6,6-d8), 2.5 mM heavy isotope-labeled [¹³C₆, ¹⁵N₂]Z-lysine (H, +8 daltons; **Supplementary Fig. 19**) or both DAP (L) and D-lysine (H). For experiments that maintained the label (all human-human cocultures), cells were initially grown for at least ten cellular doublings (ten or more days) in their respective precursors: DDC-expressing cells in DAP (L) and lyr-expressing cells in D-lysine (H). Populations were then combined in 10 mM DAP (L) and 1 mM D-lysine (H) and grown together for 5 d (approximately four cellular doublings) in coculture. Cocultures were seeded at ratios in which an equal number of cells were expected at the end of the experiment. All cell lines were passaged 1:10–1:15 at 95% confluence.

mRNA microarray expression profiling. Cells were seeded at equal densities into L-arginine and 10% dialyzed FBS-containing SILAC medium supplemented with 798 µM L-lysine, 798 µM L-lysine (M), 4 mM D-lysine HCl or 10 mM DAP. After 72 h, cells were washed, trypsinized, pelleted and frozen at –80 °C. RNA was extracted using the RNeasy mini kit (Qiagen), labeled and hybridized to Illumina Mouseref-8 or Human HT-12 microarrays. After median centering, the probe intensities for each array, moderated t-statistics and false discovery rate calculations for multiple hypothesis correction were performed using the eBayes method provided in LIMMA^{28,29}.

Drug perturbation assays. Cells were seeded in 96-well plates (2,000 cells/well) and grown to 40% confluence in SILAC medium

containing 0.798 mM L-lysine or 10 mM DAP DMEM with 10% dialyzed FBS. Cells were then inhibited with eight different drug concentrations (twofold dilution) in eight replicates. Drugs used were Stattic (STAT3 inhibitor), PI3K-IV (PI3K inhibitor), AKT-VIII (AKT inhibitor) and SL327 (MEK inhibitor). After 48-h drug treatment, cell viability was measured by Resazurin (Sigma) as described by the manufacturer. Cell viability relative to untreated cells was calculated to obtain dose-response curves.

Western blotting. Frozen cell pellets were thawed and lysed for 20 min with NP40 lysis buffer, which contained 1% Nonident P-40, 1 mM sodium orthovanadate and Complete protease inhibitors (Roche Diagnostics) in PBS. Protein concentrations were determined by the Bradford assay (BioRad) and adjusted to 1–1.5 mg/ml. Protein was then denatured in 2% SDS for 5 min at 95 °C. Approximately 20 µg of each sample was then separated by SDS-PAGE, transferred to PVDF membrane and immunoblotted using primary and secondary antibodies. Primary antibodies were obtained from Cell Signaling Technology (anti-GAPDH, #2118; anti-AKT-pS473, #4060; anti-ERK1/2-pT202/pY204, #9101; anti-p70S6K-T389, #9234S) or Abcam (anti-ATF4, #ab23760) and all experiments were carried out at a dilution of 1:1,000. The secondary HRP-conjugated anti-rabbit antibody was obtained from Cell Signaling Technology (#7074) and was used at a dilution of 1:4,000. Chemoluminescence visualization was performed on Kodak or HyBlotCL films and films were scanned by a micro-TEK scanner at 600 d.p.i. in grayscale. To test for protein loading, membranes were stripped and reprobed with anti-GAPDH (Cell Signaling Technology, #2118) at a dilution of 1:1,000.

Gene set enrichment analysis. Data from Illumina Mouseref-8 or Human HT-12 microarray experiments were analyzed with the Gene set enrichment analysis (GSEA) software package version 2.07 (ref. 30). The KEGG pathway (v2.5) was selected as gene sets database and the experimental treatment conditions (L-lysine versus precursor or L-lysine versus starved) were used as phenotype vector. Default input parameters were applied. Default significance levels were used to report perturbed KEGG pathway ontologies (nominal $P < 0.01$ and FDR < 0.25).

Mass spectrometry sample preparation. To collect cell lysate, cells were trypsinized, resuspended in SILAC DMEM and washed three times in ice-cold PBS, and cell pellets were frozen at –80 °C. For FACS samples, cocultures of GFP⁺ and mCherry⁺ cells were trypsinized, washed and resuspended in PBS with 20% medium (2% FBS) to a concentration of ~2 × 10⁷ cells/ml. Cells were then sorted into single GFP⁺ and mCherry⁺ populations on a MoFlo cell sorter (Dako) and washed twice with ice-cold PBS, and cell pellets were stored at –80 °C for further analysis. A small aliquot of each sorted population was immediately reanalyzed to determine purity. For cultured medium samples, cells were washed three times with PBS and supplied with serum-free SILAC DMEM 16 h before supernatant sample collection. Medium was collected, filtered with a 0.22 µm filter, and proteins were concentrated to ~1 mg/ml using a 3 kDa Amicon Ultra Centrifuge filter (Millipore) as described by the manufacturer.

Protein extraction and digestion. Cell pellets were resuspended with denaturation buffer (6 M urea and 2 M thiourea in 10 mM

Tris), 1 µl of benzonase was added, followed by incubation for 10 min at room temperature. Cellular debris was removed by centrifugation at 4,000g for 30 min. For the supernatant samples, the secreted proteins were precipitated by chloroform-methanol extraction. Protein concentration was assessed by the Bradford assay (Bio-Rad). Crude protein extracts were subjected to either in-gel or in-solution digestion. For the in-gel digestion, protein extracts were cleaned on a 10 cm, 4–12% gradient SDS-PAGE gel (Novex). The resulting lane was cut from the gel and subjected to in-gel digestion with trypsin and/or LysC as described previously³¹. Upon gel extraction, peptides were cleaned using Stage-tips and analyzed by nano-LC-MS/MS. For in-solution digestion, proteins from the crude extract were reduced with 1 mM dithiothreitol (DTT), alkylated with 5 mM iodoacetamide, predigested with the endoprotease LysC (Wako) for 3 h and then digested with LysC or trypsin overnight³². The resulting peptide mixture was cleaned using Stage-tips³³ and subjected to nano-LC-MS/MS without prior peptide separation.

LC-MS/MS analysis. All samples were analyzed by online nano-flow LC-MS/MS as previously described³⁴ with a few modifications. Briefly, nano-LC-MS/MS experiments were performed on an EASY-nLC system (Proxeon Biosystems) connected to an LTQ-Orbitrap XL or LTQ-Orbitrap Elite (Thermo Scientific) through a nanoelectrospray ion source. Peptides were auto-sampled directly onto the 15-cm-long 75-mm-inner-diameter analytical column packed with reversed-phase C18 ReproSil AQUA-Pur 3 mm particles at a flow rate of 500 nl/min. The flow rate was decreased to 250 nl/min after loading, and the peptides were separated with a segmented linear gradient of acetonitrile from 5–50% in 0.5% acetic acid for either 100 min, 150 min or 240 min. Eluted peptides from the column were directly electrosprayed into the mass spectrometer. For the LTQ-Orbitrap XL analyses, the instrument was operated in positive ion mode, with the following acquisition cycle: a full scan recorded in the orbitrap analyzer at resolution R of 60,000 full width at half maximum (FWHM) was followed by MS/MS collision-induced dissociation (CID) of the top ten most intense peptide ions in the LTQ analyzer. The total acquisition time was either 150 min or 240 min. For LTQ-Orbitrap Elite data acquisition, the instrument was operated in the positive-ion mode, with the following acquisition cycle: a full scan recorded in the orbitrap analyzer at resolution R of 120,000 FWHM was followed by MS/MS (CID rapid scan rate) of the 20 most intense peptide ions in the LTQ analyzer. The total acquisition time was either 100 min or 240 min depending on the method of sample preparation. Monoenzyme coculture samples were measured on the LTQ-Orbitrap XL with slight modifications: a full scan recorded in the orbitrap analyzer at resolution R of 60,000 FWHM was followed by MS/MS (CID) of the top five most intense peptide ions, with a total acquisition time of 95 min.

Processing of mass spectrometry data. The MaxQuant software package with the Andromeda search engine was used to identify and quantify proteins in cellular lysates and media^{19,35}. MaxQuant version 1.2.2.5 was used to analyze all samples except for comparing protein intensities across differentially labeled samples (**Supplementary Fig. 8**), where version 1.3.0.5 was required to select multiple groups with different labels. Mouse and human IPI protein databases (both version 3.84,

<ftp://ftp.ebi.ac.uk/pub/databases/IPI>) plus common contaminants and CTAP transgenes were used. With the exception of 'second peptides', which was deselected, default parameters were selected. Detection and quantitation of L-lysine containing peptides was specified as light L-lysine, medium L-lysine, and heavy L-lysine. For L-lysine derived from precursors DAP, z-lysine (H), and d-lysine (H), variable labels were specified as light L-lysine, heavy L-lysine (Lys8), and a custom modification (8 deuterium atoms for L-lysine), respectively. Only L-lysine-containing peptides were used for label quantitation. Default parameters in MaxQuant were used; MS/MS mass tolerance of 20 p.p.m., 1% FDR for peptide and protein identification, Re-quantify on, carbamethylation of cysteine (fixed modification), and oxidation of methionine and N-terminal protein acetylation (variable modifications).

Peptide and protein statistics (for example, sequences, H/L ratios, intensities) were extracted from MaxQuant exported peptides.txt and proteingroups.txt, respectively. Entries that MaxQuant classified as contaminants were removed. Unless otherwise stated, no other filters or normalizations were applied to the H/L ratios. Peptides were determined to be species-specific if they appeared exclusively in one of the human or mouse IPI protein databases. For the species-specific sequence determination an exact peptide sequence match to the protein database was required, except isoleucine (Ile) and leucine (Leu) were used interchangeably. Percentage of heavy label was calculated from the H/L ratio (r) as $= 100 \times r/(r + 1)$. All peptide data presented in this manuscript were filtered for human and mouse specific sequences, irrespective of whether the samples were from monocultures or cocultures. No organism-specific filtering was performed for any protein data presented. At least two peptides (razor and unique) were required for protein identification. For displaying histograms of H/L ratios, 50 bins were used and a nonparametric kernel-smoothing function was applied to fit the distribution of the histogram. All peptide and protein

lists with associated H/L ratios and intensities are provided in **Supplementary Data**.

Storage of primary data. Raw LC-MS/MS data are available on the CTAP website (<http://www.ctap.ms/>). Oligonucleotide sequences for inserts are available in **Supplementary Note 1**.

23. Kuan, Y.-C. *et al.* Biochemical characterization of a novel lysine racemase from *Proteus mirabilis* BCRC10725. *Process Biochem.* **46**, 1914–1920 (2011).
24. Nanduri, V., Goldberg, S., Johnston, R. & Patel, R. Cloning and expression of a novel enantioselective N-carbobenzyloxy-cleaving enzyme. *Enzyme Microb. Technol.* **34**, 304–312 (2004).
25. Mavrakis, K.J. *et al.* Tumorigenic activity and therapeutic inhibition of Rheb GTPase. *Genes Dev.* **22**, 2178–2188 (2008).
26. Swift, S., Lorens, J., Achacoso, P. & Nolan, G.P. Rapid production of retroviruses for efficient gene delivery to mammalian cells using 293T cell-based systems. *Curr. Protoc. Immunol.* **10**, 10.17C (2001).
27. Szymczak, A.L. *et al.* Correction of multi-gene deficiency *in vivo* using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat. Biotechnol.* **22**, 589–594 (2004).
28. Smyth, G. Limma: linear models for microarray data. in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* (eds. Gentleman, R., Carey, V., Dudoit, S., Irizarry, R. & Huber, W.) 397–420 (Springer, New York, 2005).
29. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
30. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
31. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2006).
32. Macek, B. *et al.* Phosphorylation of the human full-length protein kinase Ciota. *J. Proteome Res.* **7**, 2928–2935 (2008).
33. Ishihama, Y., Rappaport, J. & Mann, M. Modular stop and go extraction tips with stacked disks for parallel and multidimensional Peptide fractionation in proteomics. *J. Proteome Res.* **5**, 988–994 (2006).
34. Olsen, J.V. *et al.* Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127**, 635–648 (2006).
35. Cox, J. *et al.* Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).