SILAC-based Quantitative Proteome Analysis with PEAKS

Wen Zhang, Ziaur Rahman, Yifang Chen, Lei Xin, and Baozhen Shan
Bioinformatics Solutions Inc., Waterloo, Canada

Aims

To provide an overview of SILAC (Stable Isotope Labeling with Amino Acids in Cell Culture)-based data analysis in PEAKS Q.

Summary:

PEAKS supports complex SILAC data analysis and quantitative PTM profiling of individual proteins and all identified proteins. Published datasets of SILAC-based comparative quantification of global proteomes and phosphoproteomes were used to demonstrate SILAC data analysis and PTM Profile function in PEAKS Studio.

Benefits

- Accurate and sensitive detection and association of SILAC feature pairs
- Transfer IDs between associated SILAC pairs and different MS runs
- Flexible experimental design and statistical tools to facilitate data analysis

2017. Sep. Revision

Introduction

➢ SILAC and Super-SILAC

SILAC (stable isotope labeling by amino acids in cell culture) has become one of the most popular labeling techniques for mass spectrometry (MS)-based quantitative proteomics [1]. In this metabolic labeling strategy, differential isotope labeled samples (proteins/peptides) are combined early in the experimental procedure and analyzed together by LC -MS/MS. Therefore, the variation introduced from sample processing is minimized. Since stable-isotope labeled peptides have almost the same physicochemical properties as their natural counterparts, the same peptides with differential labeling co-elute from the liquid chromatography (LC) column and their amounts can be accurately quantified relative to each other.

An extension of this labeling strategy is named super-SILAC [2, 3], where labeled samples can be produced separately and spiked into each of the experimental samples that are not amenable to metabolic labeling, e.g. human tissues. This pool of spiked-in heavy proteins is used as an internal standard for quantification.

PEAKS Q supports analyses of SILAC and super-SILAC types of data.

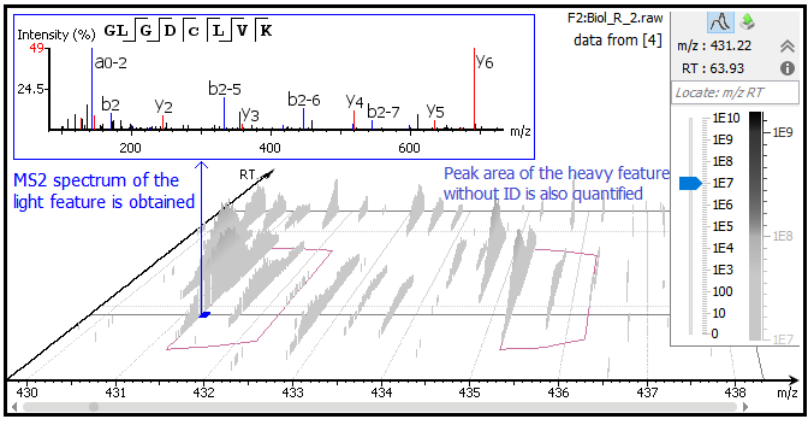
➢ PTM

Protein post-translational modifications (PTMs) play critical roles in diverse biological processes. Identification and characterization of PTMs are essential for a comprehensive understanding of cellular biology and human diseases. Recently, MS-based isotope labeling strategies such as SILAC, TMT and iTRAQ have gained popularity in quantitative proteomics, which have successfully been applied to identification and quantification of PTMs. A new PTM Profiling function has been implemented in the PEAKS Studio software for PTM analysis and direct visualization of results.

SILAC Quantification Algorithms in PEAKS Q

➢ ID transfer between associated SILAC pairs

PEAKS Q detects and associates 2 or 3-plex SILAC feature pairs that have the same charge, similar MS1 peak area correlation over retention time, expected mass shifts caused by labeling and fall within certain mass errors. If an identification is obtained from one of the



3 feature vectors

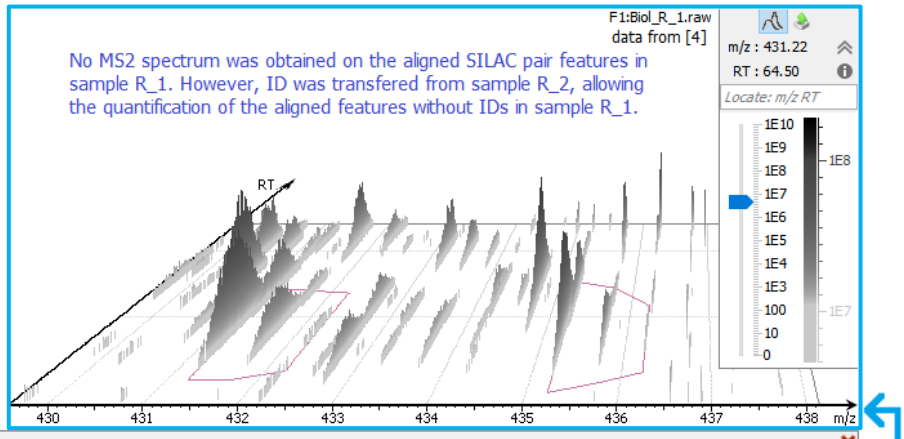
	Sample	Fraction	-10lgP	Quality Score	Light_Area	Heavy_Area	Light_Ratio	Heavy_Ratio	Id Count	m/z	z	ppm	RT	AScore
1	R_1	F 1	20.18	60.60	2.47E8	2.42E8	1.00	0.98	0	431.2292	2	1.5	63.93	
2	R_2	F 2	20.18	59.12	2.75E8	2.62E8	1.00	0.95	1	431.2292	2	1.5	63.93	
3	R_2_1	F 3	25.06	63.84	1.89E8	1.92E8	1.00	1.02	2	431.2288	2	0.5	64.30	

Peptide	-10lgP	Quality Score	Heavy : Ratio Profile	Group 1 : Light_Ratio	Group 1 : Heavy_Ratio	#Vector	Accession
GLGDC(+57.02)LVK(*)	25.06	61.19		1.00	0.98	3	P12236|ADT3_HUMAN:P05141|ADT2_HUMAN

labeled states, then the whole SILAC pair feature can be quantified and used for peptide and protein ratio calculations. For example, the light form of peptide GLGDCLVK was fragmented and identified in sample R_2 [4]. Although no MS2 spectrum was obtained from the K8 labeled heavy counterpart, the SILAC pair could still be quantified in PEAKS Q as highlighted in the feature vectors table. “Id Count” indicates the number of MS2 identified for each SILAC pair.

➢ ID transfer between different MS runs after alignment

Retention times of different LC-MS runs are first aligned, then the MS/MS and ID can be matched from another run by aligning features within tight mass ranges and retention times, allowing quantification of SILAC pairs without any ID.



3 feature vectors

	Sample	Fraction	-10lgP	Quality Score	Control_Area	Treated_Area	Control_Ratio	Treated_Ratio	Id Count	m/z	z	ppm	RT	AScore
1	R_1	F 1	20.18	60.60	2.47E8	2.42E8	1.00	0.98	0	431.2292	2	1.5	63.93	
2	R_2	F 2	20.18	59.12	2.75E8	2.62E8	1.00	0.95	1	431.2292	2	1.5	63.93	
3	R_2_1	F 3	25.06	63.84	1.89E8	1.92E8	1.00	1.02	2	431.2288	2	0.5	64.30	

Case Study 1: single-group SILAC-based data analysis with PEAKS

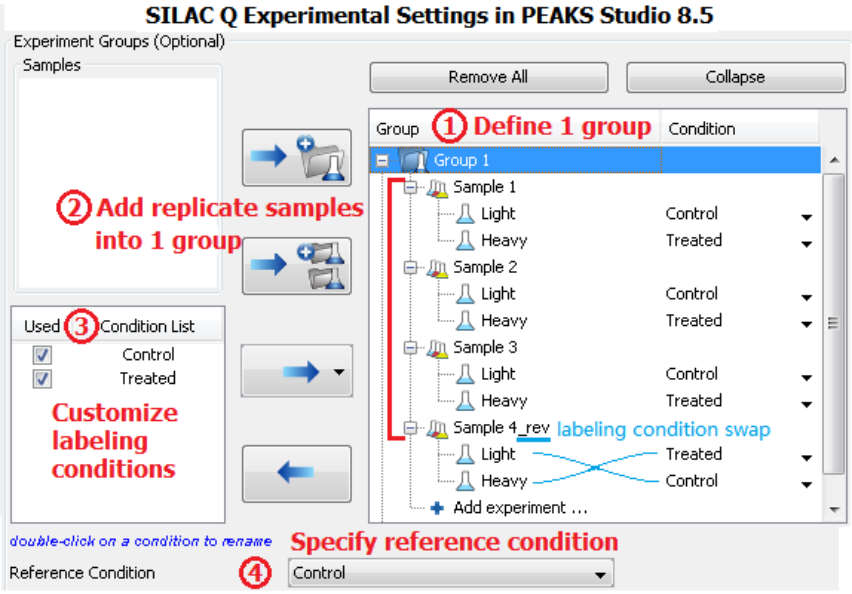
A published dataset [5] was used as an example to demonstrate the capability of SILAC-based data analysis (single group) in PEAKS Studio.

➢ Study Aims and Background

Angiotensin II (AngII) is a major effector of the renin-angiotensin system and mediates kidney disease progression. However there are no specific measures of renal AngII activity. This study aimed to define AngII-regulated proteomes, and thus to identify potential AngII activity markers in the kidney.

➢ Experimental Design

Primary human renal cells were either cultured in regular medium or in $^{13}C_6$-arginine (R6) and $^{13}C_6{}^{15}N_2$-lysine (K8) medium. After 6 passages, AngII was added to SILAC-labeled cells and proteins of the treated and control cells were extracted and mixed at 1:1 protein ratio. In total, four replicates were performed, with one containing



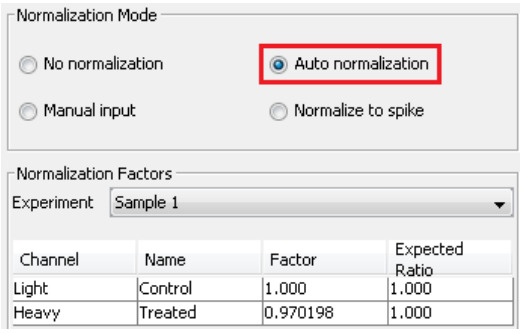
cells grown in regular medium and treated with AngII (labeling condition swapped). In the experimental settings in PEAKS Studio 8.5, the control condition was specified as the reference condition so that ratios of the treated relative to control are calculated.

➢ Data analysis

MS data was analyzed in PEAKS Studio 8.5 using a customized SILAC-2plex (R6, K8) method in PEAKS Q for quantification.

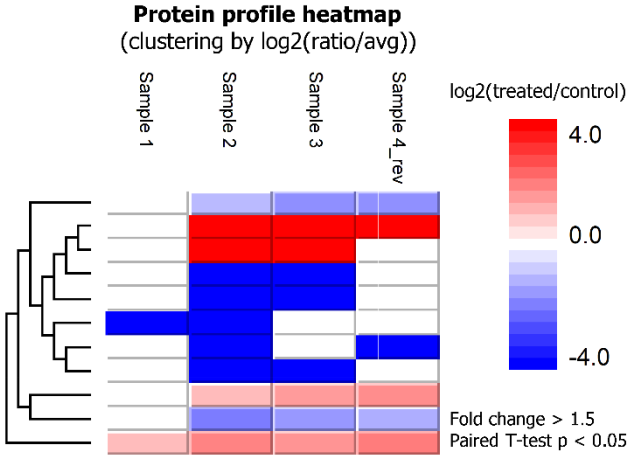
➢ Normalization of SILAC ratios in each sample

Auto normalization was first performed in PEAKS Q so that the total light and heavy intensities of each sample were equivalent since the same amount of light and heavy proteins were mixed. The adjusted normalization factor is displayed for each sample in the normalization setting window.



➢ Identification of differentially expressed proteins between groups

SILAC ratios of proteins were calculated using the median of peptide ratios (MS1 peak area in the labeling channel relative to the reference channel). Proteins that had significantly differential expressions between treated and control conditions across four replicates could be identified by applying a fold change filter of, e.g. at least 1.5, and a paired T-test p value smaller than 5%. Paired T-test is the statistical tool integrated in PEAKS Q for single-group SILAC data analysis.



➢ Conclusions

PEAKS Studio supports complex SILAC data analysis and offers paired T-test approach for significance calculation of single-group SILAC data to identify proteins with significant changes of expression levels between different conditions across replicates.

Case Study 2: multiple-group SILAC-based data analysis with PEAKS

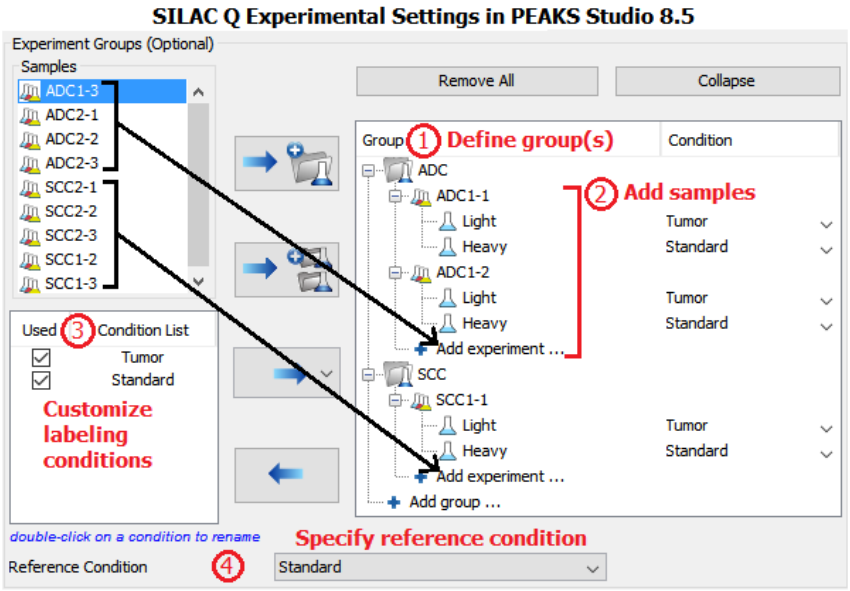
A published dataset [6] was used as an example to demonstrate the capability of SILAC-based data analysis (multiple groups) in PEAKS Studio.

➢ Study Aims and Background

Adenocarcinoma (ADC) and squamous cell carcinoma (SCC) are the two major histological subtypes of non-small cell lung cancer (NSCLC), which are associated with patient prognosis. One of the main goals of the study was to identify signature proteins that distinguished the ADC and SCC subtypes.

➢ Experimental Design

Four patient tumor-derived xenograft proteomes (two ADC and two SCC, each with 3 replicates) were quantified and compared by the super-SILAC approach. Three NSCLC cell lines were labeled with $^{13}C_6{}^{15}N_2$-lysine (K8) and $^{13}C_6{}^{15}N_4$-arginine (R10) and heavy proteins were combined and used as the super-SILAC standard. The standard was spiked into non-labeled xenograft tissues at 1:1 protein ratio. Therefore, protein expression levels in ADC and SCC xenografts were compared with reference to the super-SILAC standard.

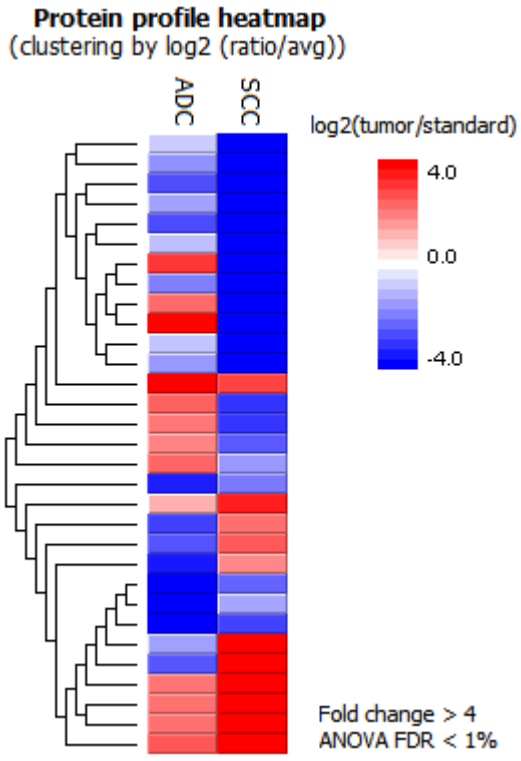


➢ Data analysis

MS data was analyzed in PEAKS Studio 8.5 using the built-in SILAC-2plex (R10, K8) method in PEAKS Q for quantification.

➢ Identification of differentially expressed proteins between groups

Proteins that had significantly differential expressions between ADC and SCC groups could be identified by applying a fold change filter of, e.g. at least 4, and an ANOVA test FDR 1% threshold. Welch's ANOVA is the statistical tool integrated in PEAKS Q for multi-group SILAC data analysis.



➢ Conclusions

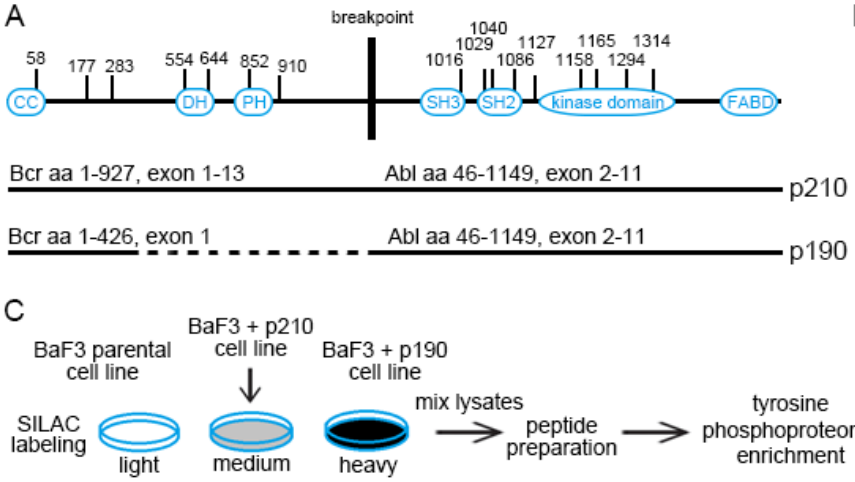
PEAKS Studio supports complex SILAC data analysis and offers ANOVA approach for significance calculation of multi-group SILAC data to identify proteins with significant changes of SILAC ratios between different groups.

Case Study 3: SILAC-based PTM profiling analysis with PEAKS

A published dataset [7] was used as an example to demonstrate the capability of SILAC-based data analysis (PTM profiling) in PEAKS Studio.

➢ Study Aims and Background

The two major isoforms of the oncogenic Bcr–Abl tyrosine kinase, p210 and p190, were suggested to contribute to different types of leukemia. One of the main goals of the study was to identify differential signaling networks of Bcr-Abl p210 and p190 kinases in leukemia cells by using quantitative proteomics approaches [7]. Bcr–Abl is a fusion protein and a constitutively active tyrosine kinase. p190 is ~25% shorter than p210 due to a lack of a DH–PH domain unit; otherwise p210 and p190 have an identical sequence and domain organization (Figure 1A).



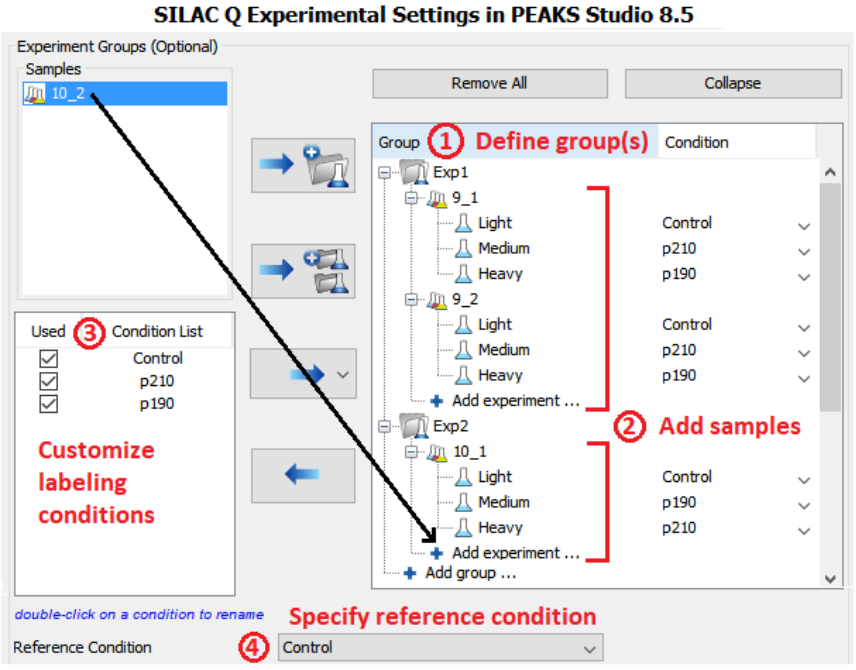
B

Experimental Design	
Exp 1: Sample 9 × 2 replicates	
Light	Control
Medium	p210
Heavy	p190
Exp 2: Sample 10 × 2 replicates	
Light	Control
Medium	p190
Heavy	p210

Figure 1. Bcr-Abl protein domain organization (A) and overview of the phosphoproteome experiments (B, C).

➢ Experimental Design

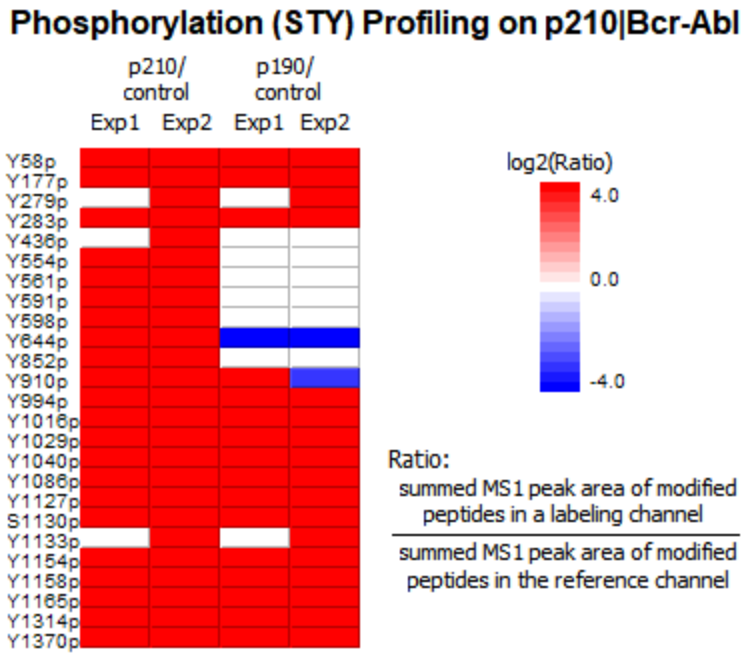
Murine BaF3 cells that express Abl endogenously were retrovirally transduced with the human Bcr–Abl p210 and p190 cDNAs. Parental (untransduced) BaF3 cells were used as a control and grown in SILAC light media (Figure 1C). Cells that expressed human p210 and p190 were labeled with SILAC medium and heavy media and swapped between 2 experiments (defined as Exp 1 and 2 in Figure 1B). Phosphotyrosine peptides were enriched and analyzed by high-resolution LC-MS/MS.



➢ Data analysis

MS data was analyzed in PEAKS Studio 8.5 using the built-in SILAC-3plex (R6, K4|R10, K8) method in PEAKS Q for quantification.

➢ The Phosphorylation Profile of p210/p190 with PEAKS PTM Profile



The phosphorylation profile of Bcr-Abl proteins were analyzed with the PTM Profile function in PEAKS. 25 phosphorylation sites of Bcr-Abl were quantified (Ascore [8] > 20). As expected p210 and p190 samples showed higher phosphorylation signals compared to controls. For each confident PTM site, MS1 peak area of modified peptides are summed for each labeling channel and ratios relative to the reference channel are calculated automatically and displayed in the heatmap, enabling easy comparative quantitation of PTM abundances between samples and across groups. Furthermore, PTM quantification results can be normalized by protein expression levels automatically.

➢ Conclusions

PEAKS Studio supports complex SILAC data analysis and quantitative PTM profiling analysis.

References

[1] Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., *et al.*, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & cellular proteomics : MCP* 2002, *1*, 376-386.
[2] Neubert, T. A., Tempst, P., Super-SILAC for tumors and tissues. *Nature methods* 2010, *7*, 361-362.
[3] Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J. R., Mann, M., Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nature methods* 2010, *7*, 383-385.
[4] Liberski, A. R., Al-Noubi, M. N., Rahman, Z. H., Halabi, N. M., *et al.*, Adaptation of a commonly used, chemically defined medium for human embryonic stem cells to stable isotope labeling with amino acids in cell culture. *Journal of proteome research* 2013, *12*, 3233-3245.
[5] Konvalinka, A., Zhou, J., Dimitromanolakis, A., Drabovich, A. P., *et al.*, Determination of an angiotensin II-regulated proteome in primary human kidney cells by stable isotope labeling of amino acids in cell culture (SILAC). *The Journal of biological chemistry* 2013, *288*, 24834-24847.
[6] Zhang, W., Wei, Y., Ignatchenko, V., Li, L., *et al.*, Proteomic profiles of human lung adeno and squamous cell carcinoma using super-SILAC and label-free quantification approaches. *Proteomics* 2014, *14*, 795-803.
[7] Reckel, S., Hamelin, R., Georgeon, S., Armand, F., *et al.*, Differential signaling networks of Bcr-Abl p210 and p190 kinases in leukemia cells defined by functional proteomics. *Leukemia* 2017.
[8] Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J., Gygi, S. P., A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nature biotechnology* 2006, *24*, 1285-1292.