Lawrence Berkeley National Laboratory

LBL Publications

Title

Applications of targeted proteomics in metabolic engineering: advances and opportunities

Permalink

https://escholarship.org/uc/item/09j1r0bd

Authors

Yunus, Ian S

Lee, Taek Soon

Publication Date

2022-06-01

DOI

10.1016/j.copbio.2022.102709

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Applications of targeted proteomics in metabolic engineering: advances and opportunities

Ian S. Yunus[1,2] and Taek Soon Lee[1,2,*]

[1]Joint BioEnergy Institute, 5885 Hollis Street, Emeryville, CA, 94608, USA
[2]Biological Systems & Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA

*Corresponding author: tslee@lbl.gov

Abstract

Optimization of metabolically engineered organisms requires good understanding of producing balanced level of pathway proteins. Targeted proteomics via selected-reaction monitoring (SRM) has been increasingly used in metabolic engineering research to detect and quantify sets of proteins with high selectivity, multiplexity, and reproducibility. In combination with metabolomics and other omics tools, targeted proteomics has helped optimize the production of many bio-based chemicals in various metabolic engineering cell factories. In this review, we present recent applications of targeted proteomics in metabolic engineering research and highlight several successful studies of targeted proteomics in boosting production of commodity and high value chemicals. Additionally, we also discuss challenges and limitations of current targeted proteomics and map opportunities for future research.

Highlights:

- Targeted proteomics is now a routine tool to verify protein expression levels
- Targeted proteomics enables multiplex quantification of selected proteins expression
- Targeted proteomics can be used to identify metabolic pathway bottleneck
- Altered native protein levels in metabolic engineering can be measured by proteomics
- Targeted proteomics supports genome-scale metabolic model and flux balance analysis

Introduction

Metabolic engineering plays an important role in our quest for building a future bioeconomy, including generating renewable fuels [1,2] and biomanufacturing a large variety of chemicals [3–5] from sustainable resources. However, transforming organisms into efficient cell factories that produce industrially-relevant titers for the large-scale production of these compounds has been challenging [6]. Maximizing gene expression in the biosynthetic pathway, though may be important, is not a universal solution to achieving high titers as excessive heterologous gene expression can cause significant burden to the cell, thereby affecting cell fitness and lowering product titers. In multi-gene biosynthetic pathways, it is also typically necessary to tune protein expression level to maximize metabolic flux towards product of interest and minimize accumulation of by-products that might be toxic to the cells [7]. Careful monitoring and quantification of protein expression levels are important keys to optimize product titers.

For more than 40 years, the nitrocellulose-based Western blot (immunoblot) analysis has been a principal method for detection and quantification of specific proteins in complex biological samples. Although this method can be conveniently used to quantify the same protein in various biological samples, assaying many different proteins simultaneously can increase development time and experimental costs. Alternatively, in the advent of significant advancements in mass-

spectrometry technology, discovery/shotgun proteomics has also been applied to characterize proteome samples. However, MS-MS acquisition in shotgun proteomics favors abundant peptides, restraining detection of low-abundance peptides. This bias towards mass analysis of the more highly abundant peptides limits the depth to which a proteome can be analyzed. To overcome the limitations of shotgun proteomics, the mass spectrometric approach of selected-reaction monitoring (SRM) can be implemented. In the last decade, targeted proteomics via SRM has emerged as the preferred technique to quantify multiple different proteins simultaneously in a sample [8]. For its importance and potential applications in various pharmaceutical and biotechnology industries, *Nature Methods* declared SRM as the "Method of the year" in 2012 [9]. Indeed, since its development, targeted proteomics via SRM has been used in various applications, some of which have been reviewed elsewhere, including applications in systems biology and translational medicine [10,11], biomedical research [12], drug efficacy biomarkers [13], identification of human pathogenic bacteria [14], and as tools for detection of foodborne pathogens [15]. In this review, we will focus specifically on the applications of targeted proteomics in metabolic engineering studies, including: *(1)* quantification of native/heterologous pathway proteins, *(2)* characterization of synthetic biology tools for metabolic engineering, *(3)* identification of pathway bottlenecks and optimization of biosynthetic pathways, *(4)* supporting analytical tools in genome-scale metabolic model (GEM) and flux balance analysis (FBA), and *(5)* other related applications to metabolic engineering.

Current methods and quantification techniques in targeted proteomics

The SRM, also called multiple reaction monitoring (MRM), was first used to detect and quantify small organic molecules in the late 1970s [16]. It was later implemented in targeted proteomics to complement discovery/shotgun proteomics for fast quantification of low abundant *S. cerevisiae* central carbon metabolism proteins with high quantitative accuracy [17]. Compared to discovery/shotgun proteomics, SRM technique provides higher sensitivity and selectivity as mass analyzers only focus on specific peptide and product ion pair, rather than scanning wide mass range window. A typical workflow of targeted proteomics via SRM is shown in Fig. 1. Among different protein quantification methods listed in Table 1, label-free relative quantification method and absolute quantification method using standard labelled synthetic peptides (e.g., QconCAT and AQUA) are two commonly used methods in metabolic engineering studies.

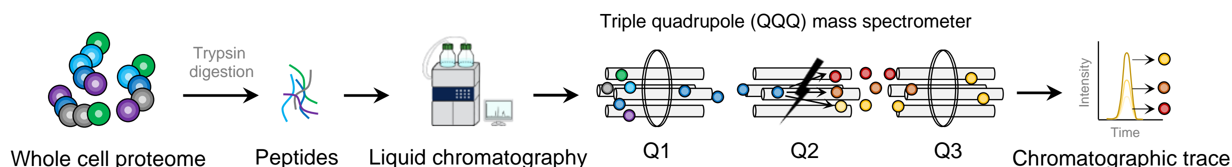

Fig 1. A typical workflow of targeted proteomics via SRM starts with cell lysis and protein extraction of biological samples followed by digestion of the proteins into peptides by a protease (e.g., trypsin). Peptides generated by trypsin cleavage are then run on a liquid chromatography column coupled to a triple quadrupole (QQQ) mass spectrometer. SRM uses the unique capability of triple quadrupole mass spectrometers to specifically filter selected peptide within a mass range centered around the selected peptide in the first mass analyzer (Q1), which is fragmented by collision-induced dissociation (CID) in the second quadrupole (Q2) to generate fragment ions. The generated fragment ions are then transferred to the third quadrupole (Q3), where only a selected m/z ion can pass, resulting in a chromatographic trace with retention time and signal intensity as coordinates [8,18]. The peptide-fragment ion pair is known as transition and the area under the chromatographic peaks for each transition is a measure of the amount/concentration of the representative protein in the sample.

Table 1. Quantification methods used in targeted proteomics

Quantification	Method	Example	Ref
Relative	Label-free strategies	Endogenous reference proteins	[19]
Relative/absolute	Metabolic stable-isotope labeling	^{13}C labeling	[20]
		^{15}N labeling	[21]
		SILAC	[22]
		SILAM	[23]
	Enzymatic stable-isotope labeling	$H_2O/H_2{}^{18}O$ labeling	[24]
	Chemical stable-isotope labeling	Dimethyl labeling	[25]
		iTRAQ	[26]
		mTRAQ	
		TMT	
		ICAT	[27]
		Photocleavable ICAT	
		ICPL	
	Standard labeled synthetic peptides	AQUA	[28–30]
		PSAQ	[31]
		QconCAT	[32–35]

SILAC, stable-isotope labeling by amino acids in cell culture; SILAM, stable-isotope labeling by amino acids in mammals; iTRAQ, isobaric tags for relative and absolute quantification; mTRAQ, mass differential tags for relative and absolute quantification; TMT, tandem mass tags; ICAT, isotope-coded affinity tags; ICPL, isotope-coded protein labels; AQUA, standard labelled synthetic peptides for Absolute QUAntification, QconCAT, recombinant expression of a quantification concatemer

Label-free relative quantification method

Label-free relative quantification method usually uses normalization strategy of proteotypic peptides of the targeted proteins against endogenous reference proteins/housekeeping proteins (i.e., proteins whose expressions are assumed unaffected by experimental conditions) [19]. Due to its ease of use without requiring expensive isotope-labeling, this method has been used in many metabolic engineering studies (Table 2). This normalization strategy also allows compensation for analyte loss during sample preparation or variability during LC-MS measurement. Although this method in many cases is sufficient to obtain relative quantification of overexpressed proteins in biosynthetic pathways between samples, recent studies indicate that stable expression of housekeeping proteins should not always be taken for granted [36]. If housekeeping proteins are used for normalization, one should validate their expression profiles are stable across the tested experimental conditions.

Stable-isotope synthetic peptides for absolute quantification method

To have a more precise quantification, several absolute quantitative proteomics approaches have been developed, utilizing different stable-isotope-labeled internal standards (IS), such as synthetic peptides (AQUA) [28–30], quantification concatemers (QconCATs) [32–35], and full-length protein standards (PSAQ) [31]. To perform AQUA, isotope-labeled synthetic peptides are added to digested protein samples as the IS, followed by peptide extraction and MS analysis.

128 Protein quantity is determined according to the ratio of peak intensities of unlabeled natural
129 peptides to their heavy isotope-labeled counterparts. Another commonly used absolute protein
130 quantification method is QconCATs which are obtained via the expression of artificial QconCATs
131 genes in *E. coli* grown in a heavy isotope enriched medium. QconCATs are added to protein
132 samples before digestion, and the digested concatenated peptides serve as IS for different
133 proteins. An alternative approach to AQUA and QconCAT is PSAQ, where full-length isotope
134 labeled proteins were used as internal standard for absolute quantification. When added to protein
135 samples before digestion, PSAQ corrects for protein losses that could occur during sample
136 preparation and LC-MS analysis. Although PSAQ is an ideal choice for absolute quantification of
137 proteins, this method is seldom used in quantification of a large number of proteins in metabolic
138 engineering studies due to the expensive nature of PSAQ proteins as full-length of isotope labeled
139 proteins must be synthesized.
140
141 Applications of targeted proteomics in metabolic engineering
142
143 One of the earliest studies to demonstrate the implementation of targeted proteomics in microbial
144 metabolic engineering was published in 2011. The study described the use of targeted proteomics
145 to identify the bottleneck enzymes in the biosynthetic pathway of amorphadiene, an antimalarial
146 drug precursor, in *E. coli* [37]. Since then, targeted proteomics has been implemented to assist
147 production of various compounds, including L-tyrosine [38], isoprenol [39,40], bisabolene [40,41],
148 limonene [40,41], isoprene [42••], farnesene [43], lycopene [44], alkane and alkene [45], n-butanol
149 [46], ethanol [47•,48••], free fatty acids and fatty alcohols [49], fatty acid methyl esters [50•],
150 spinosyn [51], *ent*-kaurene [52•], benzyl glucosinolate [53••], and many other compounds (Table
151 2). The production host is also not only limited to *E. coli*. Other organisms, such as *Streptomyces*
152 [51], *Clostridium cellulolyticum* [54•], *Corynebacterium glutamicum* [55••], *Rhodosporidium*
153 *toruloides* [52•], *Chlamydomonas reinhardtii* [56,57], a cyanobacterium *Synechocystis* sp. PCC
154 6803 [47–50], and many others have been used as production hosts. This section will describe
155 the applications of targeted proteomics in metabolic engineering studies in more details (Fig. 2).
156

Table 2. Applications of targeted proteomics in metabolic engineering

Application	Pathway	Organism host	Protein	Quantification method	Target compound (titer)	Ref
Multiplexing quantification of native/heterologous pathway proteins	Central carbon metabolism	*C. glutamicum*	10	QconCAT	None	[58]
	Central carbon metabolism	*S. cerevisiae*	137	SPPS	None	[17]
	Central carbon, amino acid metabolism	*S. cerevisiae*	137	13C labeling	None	[59]
	Central carbon metabolism	*Synechocystis* sp. PCC 6803	112	15N labeling	None	[60]
	Glycolytic pathway	*S. cerevisiae*	27	QconCAT	None	[63]
	Calvin cycle, photosynthetic apparatus, starch synthesis, glycolysis, TCA cycle, carbon concentrating mechanisms (CCM)	*Chlamydomonas reinhardtii*	88	SPPS	None	[57]
	Terpene pathway	*Picea abies* (bark)	16	SPPS	None	[64]
	Carbon and nitrogen metabolism pathway	*Medicago truncatula*	3	SPPS	None	[65]
	Fruit ripening pathway	*Fragaria* x *ananassa*	101	Label-free	None	[66]
	Polyketide pathway	*Streptomyces*	23	Label-free	Spinosyn (1.46 mg/L)	[51]
	Sucrose metabolism	*Arabidopsis thaliana*	1	SPPS	None	[65]
	Central carbon, amino acid metabolism	*S. cerevisiae*	135	13C labeling	None	[20]
	Ribosome, glycolytic pathway	*S. cerevisiae*	78	QconCAT	None	[71]
	Central carbon metabolism	*P. putida* KT2440	132	13C/15N labeling	None	[73•]
	Central carbon metabolism	*E. coli*	22	PSAQ	NADPH	[76]
	Fatty acid synthesis pathway	*E. coli*	12	PSAQ	Free fatty acids (4000 mg/L)	[77]
	Isoprenoid pathway	*C. cellulolyticum*	1	Label-free	Isoprene (20 μM)	[54•]
	Wood–Ljungdahl pathway	*Clostridium ljungdahlii*	7	Label-free	Acetone, isopropanol	[79•]
	Methylerythritol 4-phosphate (MEP) pathway	*E. coli*	1	Label-free	Isoprene (1.2 nM/OD600.min)	[42••]
	Mevalonate pathway	*R. toruloides*	2	Label-free	*ent*-kaurene (1400 mg/L)	[52•]
	Fatty acid synthesis pathway	*Synechocystis* sp. PCC 6803	4	Label-free	Fatty acid methyl esters (120 mg/L)	[50•]
	Ethanol	*Synechocystis* sp. PCC 6803	5	Label-free	Ethanol (200 mg/L/OD730)	[47•]
	Polyketide pathway	*E. coli*	2	Label-free	Alkane and alkene (140 mg/L)	[45]
	Mevalonate pathway	*E. coli*	1	Label-free	Isoprenol (2230 mg/L)	[80]
	Mevalonate pathway	*E. coli*	10	Label-free	Amorphadiene (3500 mg/L)	[81]
	Mevalonate pathway	*E. coli*	5	Label-free	Amorphadiene (700 mg/L)	[82]

	Mevalonate pathway	*E. coli*	9	Label-free	Amorphadiene (1600 mg/L)	[7]
Characterization of synthetic biology tools	Red fluorescent protein	*E. coli*	1	Label-free	None	[37]
	Ethanol pathway	*Synechocystis* sp. PCC 6803	5	Label-free	Ethanol (200/mg/L/OD730)	[47•]
Identification of pathway bottlenecks	Mevalonate, tyrosine pathways	*E. coli*	24	Label-free	Tyrosine (250 mg/L)	[38]
	Terpene pathway	*E. coli*	9	Label-free	amorpha-4,11-diene (500 mg/L)	[37]
	Tyr metabolic pathway	*E. coli*	11	Label-free	L-tyrosine (>2000 mg/L)	[78]
	Clostridial n-butanol pathway	*C. cellulolyticum*	5	Q-Tag fusions	n-butanol (120 mg/L)	[46]
	Mevalonate pathway	E. coli	7	Label-free	Isoprenol (1500 mg/L)	[39•]
	Fatty acid synthesis pathway	*Synechocystis* sp. PCC 6803	1	Label-free	Fatty alcohols (100 mg/L)	[49]
	Mevalonate pathway	*C. glutamicum*	5	Label-free	Isoprenol (1250 mg/L)	[55••]
	Glucosinolate pathway	*E. coli*	10	15N labeling	Benzyl glucosinolate (8.3 mg/L)	[53••]
	Mevalonate pathway	*E. coli*	9	Label-free	Farnesene (1100 mg/L)	[43]
	Mevalonate pathway	*E. coli*	10	Label-free	Lycopene (1440 mg/L)	[44]
Understanding of pathway regulations	Fatty acid synthesis pathway	*E. coli*	23	Label-free	Free fatty acids (5200 mg/L)	[83]
	Flavonoid/anthocyanin pathway	*Fragaria* x *ananassa*	21	Label-free	None	[72]
Supporting analytical tools in computational/ mathematical model	Central carbon, amino acid metabolism	*S. cerevisiae*	228	15N labeling	None	[21]
	Glycolytic pathway	*S. cerevisiae*	27	QconCAT	None	[74]
	Mevalonate pathway	*E. coli*	17	QconCAT	Amorphadiene	[75]
	Mevalonate pathway	*E. coli*	9	Label-free	Bisabolene (1150 mg/L) and limonene (600 mg/L)	[41•]
Supporting analytical tools in multiomics analysis	Ethanol	*Synechocystis* sp. PCC 6803	99	Label-free	Ethanol (118 mg/L)	[48••]
	Mevalonate pathway	*E. coli*	>20	Label-free	Isoprenol (300 mg/L/OD600)	[40]
Strain characterization upon environmental perturbation	Central metabolic pathways	*Synechocystis* sp. PCC 6803	106	Label-free	None	[61]
	Photosynthetic apparatus	*Synechocystis* sp. PCC 6803	244	15N labeling	None	[62•]
	Central metabolic pathways	*E. coli*	>400	QconCAT	None	[67•]
	Central carbon metabolism	*B. subtilis*	41	QconCAT	None	[68]
	Central carbon metabolism	*C. glutamicum*	19	15N labeling	None	[69]
	Organohalide respiration	*Dehalococcoides mccartyi*	10	SPPS	None	[70]

159

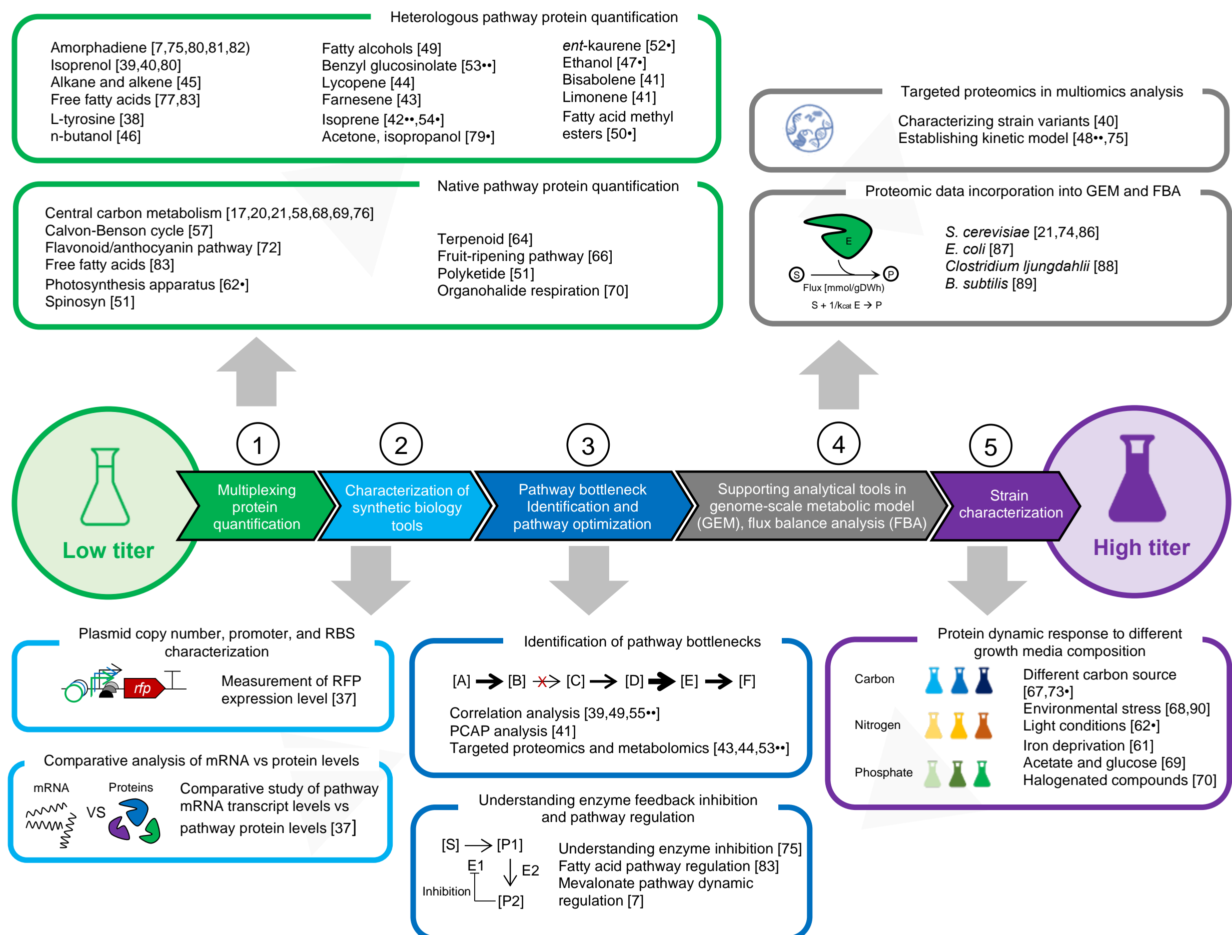

Figure 2. Application of targeted proteomics in metabolic engineering. PCAP, principal component analysis of proteomics; RFP, red fluorescent protein

Targeted proteomics for multiplexing protein quantification

One of the key advantages of MS-based approach as opposed to traditional western blotting is the capacity of MS-based proteomics to provide high throughput and multiplex protein quantification. Traditionally, western blotting is used as a simple yet powerful method to provide a yes or no answer about the presence of recombinant protein expression in a complex protein sample. However, western blotting can only be carried out if a primary antibody against the protein of interest is available. Alternatively, in the advance of fluorescence-activated cell sorting (FACS) technology, a protein of interest can be fused to a fluorescent protein (e.g., green fluorescence protein (GFP)) and the measurement of fluorescence level is used as a proxy to determine the expression level of the recombinant protein [84]. Both methods are, however, constrained by the limited number of available affinity tags and fluorescent proteins. Additionally, they require an addition of a protein tag incorporated to either the N- or C-terminus of a protein, which is not only complex to perform but can also influence the protein folding, expression, and activity. In metabolic engineering studies, especially those involving multi-gene pathways, measurement of protein expression levels using targeted proteomics is the preferred solution. Targeted proteomics has been used to confirm the expression of approximately 23 genes in the spinosyn biosynthetic gene cluster in three different *Streptomyces* species [51]. Gaida et al., introduced n-butanol biosynthesis pathway comprising five genes from *C. acetobutylicum* to *C. cellulolyticum* and

183 measured the expression of functional enzymes using targeted proteomics [46]. More recently,
184 Phillips et al. performed conjugal transfer of heterologous acetone biosynthesis pathway to an
185 industrial promising syngas-fermenting organism, *C. ljungdahlii,* and confirmed the protein
186 expression level by targeted proteomics [79•]. Targeted proteomics can also be used to monitor
187 the presence of specific proteins in a cell-free system [85•]. By using microflow liquid
188 chromatography-selected reaction monitoring (LC-SRM), Gao et al. accelerated the time needed
189 to quantify more than 100 proteins from *P. putida* KT2440 by 3-fold [73•]. Data from targeted
190 proteomics can provide useful insights on cellular metabolism and for further metabolic
191 engineering steps to improve the product titer.
192
193 **Targeted proteomics for characterization of synthetic biology tools**
194
195 Several synthetic biology tools, such as promoters, ribosome-binding sites (RBS), reporter
196 proteins, and modular assembly of genetic parts have been developed to accelerate the genetic
197 engineering of microbial hosts. One of the simplest strategies to tune protein expression level is
198 to alter regulatory elements such as promoters or to express genes in different copy number
199 plasmids. Redding-Johanson et al. used targeted proteomics via SRM to characterize several
200 promoters for heterologous protein expression in *E. coli* in different plasmid backbones by
201 measuring the red fluorescent protein expression levels [37]. The same strategy was later used
202 to monitor relative expression levels of multiple proteins in amorphadiene biosynthesis pathway
203 (Fig. 2B). By comparing the protein expression level with the mRNA transcription data, they found
204 striking differences between the transcript and protein data. The codon-optimized version of the
205 phosphomevalonate kinase gene results in no change in the level of transcript in comparison to
206 the non-codon-optimized gene whereas the protein level increased by nearly 3-fold, highlighting
207 the importance of complementing transcript data with protein expression level. By codon-
208 optimizing both mevalonate kinase and phosphomevalonate kinase genes and expressing them
209 under a stronger promoter, the final amorphadiene titer is improved from 75 to 500 mg/L. Similar
210 strategy was used by Singh et al. [38] to characterize several inducible and constitutive promoters
211 for bioproduction of L-tyrosine in *E. coli*. By fine-tuning the expression system, they successfully
212 improved the L-tyrosine production titer from 1 to 250 mg/L. In addition, they also used targeted
213 proteomics to quantify native protein levels involved in L-tyrosine biosynthetic pathway and found
214 altered expression levels of native proteins that would otherwise be difficult to detect without
215 targeted proteomics. In a more recent study, Bartasun et al. used targeted proteomics to
216 systematically characterize a set of ribosome binding site sequences in a multigene one-operon
217 system in a cyanobacterium *Synechocystis* sp. PCC 6803 for ethanol production [47•]. They found
218 that the expression level of the first gene in an operon influences the expression level of
219 subsequent genes, which is also observed elsewhere [43].
220
221 **Targeted proteomics for pathway bottleneck identification and pathway**
222 **optimization**
223
224 Low bioproduct titers can rise from inefficient translation of pathway genes, accumulation of toxic
225 intermediates, pathway enzyme inhibition, substrate competition, pathway imbalance, and many
226 other factors. A previous study by Redding-Johanson et al. [37••] had identified that mevalonate
227 kinase and phosphomevalonate kinase were the potential bottlenecks in production of
228 amorphadiene, an antimalarial drug precursor. These data provide useful information for follow-
229 up work, reducing time-consuming engineering steps. To relieve the pathway bottleneck,
230 Nowroozi et al. [81] used combinatorial expression of amorphadiene biosynthesis pathway using
231 different RBS and carefully monitored the relative pathway protein expression levels by targeted

proteomics. After selecting the most appropriate RBS combination, they successfully relieved the pathway bottleneck, reduced accumulation of toxic metabolite intermediates ((3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) and farnesyl pyrophosphate (FPP)), and improved the growth, leading to production of more than 3,500 mg/L amorphadiene. In two other studies, Zhu et al. measured the expression levels of proteins involved in mevalonate pathway using targeted proteomics and identified isopentenyl diphosphate isomerase (Idi) as one of the key enzymes for improved farnesene and lycopene bioproduction [43,44]. Pathway bottleneck can also be pinned down by plotting the metabolite/product titer against increasing enzyme expression levels. A high level correlation from linear regression may indicate the protein being monitored as a rate-limiting enzyme [49,55••]. This strategy, however, might be a challenge when the intermediate metabolites are unstable, converted into other forms, and/or difficult to monitor. Another alternative is to measure the absolute pathway protein concentration. The data can be incorporated into enzyme kinetic-based ordinary differential equation models to predict product titers. Weaver et al. [75] specifically built targeted proteomics assay to measure the absolute amorphadiene pathway protein concentration and predict the amorphadiene formation. They found good amorphadiene titer data agreement between model and experiment. In a more recent study, Petersen et al. identified and addressed metabolic bottlenecks in production of benzyl glucosinolate, a plant secondary metabolite with diverse health benefits, by changing genetic regulatory elements and monitoring pathway protein expression levels [53••].

Targeted proteomics as supporting analytical tools in multi-omics analysis and genome-scale metabolic models

Even though quantification of protein levels and target products alone can help identify pathway bottlenecks in many metabolic engineering studies, quantification of precursor/substrate, fermentation by-products, and intermediate metabolites through metabolomics and the following analysis with aid of mathematical models are often necessary to rationally identify pathway balance, bottlenecks, and potential engineering targets. For example, George et al. [39•] constructed a series of pathway variants for bioproduction of isoprenol in *E. coli*. Aided by targeted proteomics and metabolomics, they performed correlation analysis between enzyme vs metabolite levels from each pathway variant and constructed a conceptual model of isoprenol pathway behavior for further engineering steps. With a properly balanced pathway, 1.5 g/L of isoprenol was produced at 46% theoretical yield. In another example, Alonso-Gutierrez et al. [41] measured the correlation of limonene and bisabolene product titers vs enzyme levels and used principal component analysis of proteomics (PCAP) to pinpoint specific enzymes that need to have their expression level adjusted to maximize limonene and bisabolene production in *E. coli*. More recently, Volke et al., used information from metabolomics and targeted proteomics to pinpoint isopentenyl diphosphate isomerase (Idi) and 1-deoxyxululose 5-phosphate synthase (Dxs) as major flux controlling enzymes in the methylerythritol phosphate (MEP) pathway [42••]. Nishiguchi et al. established kinetic models combining metabolic flux, metabolite concentration, and protein abundance data and identified phosphoglycerate kinase as a promising engineering target to improve pyruvate supply for ethanol production [48••]. As currently only a few genome-scale metabolic models have incorporated protein abundance data (e.g., *S. cerevisiae* [86], *E. coli* [87], *C. ljungdahlii* [88], *B. subtilis* [89]), MS-based proteomics will continue to play important roles in genome-scale metabolic model development. These examples demonstrate that in combination with mathematical model and computational biology tools, proteomic data can give meaningful information to reduce iterative trial-and-error steps in the design-build-test-learn (DBTL) cycle and bring biological engineering closer to more predictable and rational engineering processes.

283 **Targeted proteomics for protein quantification in varying environmental conditions**
284
285 Optimization of production medium is required to maximize yield. The most suitable growth
286 conditions (e.g., pH, temperature, agitation speed, aeration, etc) and medium composition (e.g.,
287 carbon and nitrogen source and ratio, phosphate, etc) must be optimized accordingly. Too often
288 than not, the success of bioproduction of commodity chemicals largely depends on the feedstock
289 used. However, most medium optimization studies only use product titer as a measure of success
290 without pursuing deeper understanding at the cellular level. In the advent of omics tools, targeted
291 proteomics can provide insight into changes in the proteome level upon environmental
292 perturbations (e.g., variation in growth medium, environmental stress) [68,90]. *P. putida* KT2440,
293 an emerging workhorse for bioproduction of various chemicals [91], is of interest for many
294 metabolic engineers as this microbe can utilize various carbon sources. Kukurugya and co-
295 workers recently investigate the metabolic phenotype enabling *P. putida* KT2440 to utilize mixed
296 substrates [92]. In combination with metabolomics and fluxomics, they unravel constitutive tuning
297 of the metabolic architecture allowing co-utilization of carbohydrate and aromatic substrates in *P.*
298 *putida* KT2440. Indeed, in a similar study by Gao and co-workers where they utilize high-
299 throughput large-scale targeted proteomics assay to quantify 132 proteins (339 peptides) in *P.*
300 *putida* KT2440 grown in various carbon sources [73•], they find *P. putida* KT2440 dynamically
301 changes their central carbon metabolism protein abundance in response to different growth
302 media. In the future, a systematic study of the effect of growth medium on metabolite production
303 might pinpoint metabolic engineers how to optimize the production medium to achieve high
304 product yield.
305
306 **Targeted proteomics in pathway discovery**
307
308 Model organisms such as *E. coli* and *S. cerevisiae* are still the most widely used organisms for
309 metabolic engineering purposes. Recently, extremophiles bacteria have gained interest as non-
310 model metabolic engineering platforms due to their fascinating lifestyle [93,94]. Although limited
311 available metabolic engineering tools make them hard to engineer, extremophilic organisms may
312 serve as object studies of non-model organisms to elucidate the molecular basis of survival or
313 adaptive response. A proteomics analysis of the hydrocarbon degrading *Oleispira antarctica* RB-
314 8 revealed a n-alkane oxidation pathway consisting of several alkane monooxygenases, alcohol
315 and aldehyde dehydrogenases, a fatty acid-CoA ligase, and a fatty acid desaturase. When grown
316 on tetradecane, these proteins were upregulated by 3 to 21-fold, shedding some light on
317 hydrocarbon degradation pathway in this organism [95]. Though obtained from shotgun
318 proteomics, such information can further be used to identify enzymes produced by extremophiles
319 that have biotechnological and commercial value (e.g., enzymes with thermal stability, higher
320 activity, or pH and solvent tolerance) [96,97]. The identified enzymes then can be heterologously
321 expressed in model hosts for industrial applications.
322

323 Perspectives

324
325 Targeted proteomics has been a useful tool in many metabolic engineering studies. Recent years
326 have seen increasing efforts to synergistically combine targeted proteomics with other systems
327 biology tools (omics, GEM, etc) to bring a more predictable and rational engineering of biology.
328 Much, however, remains to be done in order to allow rapid development of cell factories with
329 industrially relevant titers. Recently, machine learning [98] has emerged as an effective tool to
330 predict pathway optimization [99], model RBS sequence – phenotype relationship [100], and
331 generate Automated Recommendation Tool (ART) [101]. Integration of high quality and accuracy

protein expression data input from targeted proteomics and other omics tools to machine learning will become an avenue of interest.

Conflict of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

• of special interest

•• of outstanding interest

[42••] This study demonstrates the use of data from metabolomics and targeted proteomics in combination with recombineering for precise metabolic control analysis of methylerythritol 4-phosphate pathway.

[47•] This study uses targeted proteomics to carefully quantify the relative expression level of proteins in multi-gene pathways and shows the influence of operon structure on protein expression levels.

[48••] This paper combines targeted proteomics, metabolic flux analysis, and metabolomics to obtain kinetic model for rationally designing engineered microorganisms.

[50•] This study describes the use of targeted proteomics to compare the expression level of heterologously expressed proteins in different engineered *Syenchocystis* strains.

[52•] This study uses targeted proteomics to quantify the relative expression level of heterologous proteins in *ent*-kaurene biosynthesis pathway. This study also shows that increasing transcript levels does not always yield higher protein expression levels.

[53••] This paper describes the use of targeted proteomics to pinpoint the bottleneck in the benzyl glucosinolate pathway.

[54•] This study demonstrates the use of targeted proteomics to confirm the presence of heterologously expressed protein in *Clostridium cellulolyticum*.

[55••] This study identifies the potential rate-limiting enzyme in mevalonate pathway-based isoprenol production in *Corynebacterium glutamicum* using targeted proteomics.

[62•] Recent application of targeted proteomics to understand the effect of varying light conditions on photosynthetic apparatus in *Synechocystis* sp. PCC 6803

[73•] This paper describes the use of microflow liquid chromatography-mass spectrometry selected reaction monitoring for high-throughput protein quantification workflow in *P. putida* KT2440.

[79•] This study demonstrates the use of targeted proteomics to confirm the presence of heterologously expressed protein in *Clostridium ljungdahlii*.

[85•] This study shows the application of targeted proteomics to confirm the presence of expressed proteins in cell-free systems.

CRediT authorship contribution statement

Ian S. Yunus: Conceptualization, Visualization, Writing – Original draft preparation, review & editing**. Taek Soon Lee:** Conceptualization, Supervision, Writing - review & editing

380 ## Acknowledgements

381

382 This work was supported by the DOE Joint BioEnergy Institute (http://www.jbei.org), supported
383 by the US Department of Energy, Office of Science, Office of Biological and Environmental
384 Research, through contract DE-AC0205CH11231 with Lawrence Berkeley National Laboratory.
385 The United States Government retains and the publisher, by accepting the article for publication,
386 acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable,
387 worldwide license to publish or reproduce the published form of this manuscript, or allow others
388 to do so, for United States Government purposes. The views and opinions of the authors
389 expressed herein do not necessarily state or reflect those of the United States Government or
390 any agency thereof. Neither the United States Government nor any agency thereof, nor any of
391 their employees, makes any warranty, expressed or implied, or assumes any legal liability or
392 responsibility for the accuracy, completeness, or usefulness of any information, apparatus,
393 product, or process disclosed, or represents that its use would not infringe privately owned rights.

394
395

396 ## References

397

398 1. Liu Y, Cruz-Morales P, Zargar A, Belcher MS, Pang B, Englund E, Dan Q, Yin K,
399 Keasling JD: **Biofuels for a sustainable future**. *Cell* 2021, **184**:1636–1647.
400 2. Keasling J, Garcia Martin H, Lee TS, Mukhopadhyay A, Singer SW, Sundstrom E:
401 **Microbial production of advanced biofuels**. *Nat Rev Microbiol* 2021,
402 doi:10.1038/s41579-021-00577-w.
403 3. Lee SY, Kim HU, Chae TU, Cho JS, Kim JW, Shin JH, Kim DI, Ko Y-S, Jang WD, Jang
404 Y-S: **A comprehensive metabolic map for production of bio-based chemicals**. *Nat*
405 *Catal* 2019, **2**:18–33.
406 4. Lu H, Villada JC, Lee PKH: **Modular Metabolic Engineering for Biobased Chemical**
407 **Production**. *Trends Biotechnol* 2019, **37**:152–166.
408 5. Choi KR, Jang WD, Yang D, Cho JS, Park D, Lee SY: **Systems Metabolic Engineering**
409 **Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering**.
410 *Trends Biotechnol* 2019, **37**:817–837.
411 6. Nielsen J, Keasling JD: **Engineering Cellular Metabolism**. *Cell* 2016, **164**:1185–1197.
412 7. Dahl RH, Zhang F, Alonso-Gutierrez J, Baidoo E, Batth TS, Redding-Johanson AM,
413 Petzold CJ, Mukhopadhyay A, Lee TS, Adams PD, et al.: **Engineering dynamic**
414 **pathway regulation using stress-response promoters**. *Nat Biotechnol* 2013, **31**:1039–
415 1046.
416 8. Lange V, Picotti P, Domon B, Aebersold R: **Selected reaction monitoring for**
417 **quantitative proteomics: a tutorial**. *Mol Syst Biol* 2008, **4**:222.
418 9. Gillette MA, Carr SA: **Quantitative analysis of peptides and proteins in biomedicine**
419 **by targeted mass spectrometry**. *Nat Methods* 2013, **10**:28–34.
420 10. Ebhardt HA, Root A, Sander C, Aebersold R: **Applications of targeted proteomics in**
421 **systems biology and translational medicine**. *Proteomics* 2015, **15**:3193–3208.
422 11. Manes NP, Nita-Lazar A: **Application of targeted mass spectrometry in bottom-up**
423 **proteomics for systems biology research**. *J Proteomics* 2018, **189**:75–90.
424 12. Shi T, Song E, Nie S, Rodland KD, Liu T, Qian W-J, Smith RD: **Advances in targeted**
425 **proteomics and applications to biomedical research**. *Proteomics* 2016, **16**:2160–
426 2182.
427 13. Masuda T, Mori A, Ito S, Ohtsuki S: **Quantitative and targeted proteomics-based**
428 **identification and validation of drug efficacy biomarkers**. *Drug Metab Pharmacokinet*
429 2021, **36**:100361.

430 14. Saleh S, Staes A, Deborggraeve S, Gevaert K: **Targeted Proteomics for Studying**
431 **Pathogenic Bacteria**. *Proteomics* 2019, **19**:1800435.
432 15. Shevchuk O: **3.11 - Targeted Proteomics for Rapid and Sensitive Detection of**
433 **Foodborne Pathogens**. In Edited by Cifuentes ABT-CF. Elsevier; 2021:123–136.
434 16. Yost RA, Enke CG: **Selected ion fragmentation with a tandem quadrupole mass**
435 **spectrometer**. *J Am Chem Soc* 1978, **100**:2274–2275.
436 17. Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R: **Full Dynamic Range**
437 **Proteome Analysis of S. cerevisiae by Targeted Proteomics**. *Cell* 2009, **138**:795–
438 806.
439 18. Picotti P, Aebersold R: **Selected reaction monitoring–based proteomics: workflows,**
440 **potential, pitfalls and future directions**. *Nat Methods* 2012, **9**:555–566.
441 19. Bluemlein K, Ralser M: **Monitoring protein expression in whole-cell extracts by**
442 **targeted label- and standard-free LC-MS/MS**. *Nat Protoc* 2011, **6**:859–869.
443 20. Matsuda F, Kinoshita S, Nishino S, Tomita A, Shimizu H: **Targeted proteome analysis**
444 **of single-gene deletion strains of Saccharomyces cerevisiae lacking enzymes in**
445 **the central carbon metabolism**. *PLoS One* 2017, **12**:e0172742.
446 21. Costenoble R, Picotti P, Reiter L, Stallmach R, Heinemann M, Sauer U, Aebersold R:
447 **Comprehensive quantitative analysis of central carbon and amino-acid metabolism**
448 **in Saccharomyces cerevisiae under multiple conditions by targeted proteomics**.
449 *Mol Syst Biol* 2011, **7**:464.
450 22. Chen X, Wei S, Ji Y, Guo X, Yang F: **Quantitative proteomics using SILAC:**
451 **Principles, applications, and developments.** *Proteomics* 2015, **15**:3175–3192.
452 23. Rauniyar N, McClatchy DB, Yates JR: **Stable isotope labeling of mammals (SILAM)**
453 **for in vivo quantitative proteomic analysis**. *Methods* 2013, **61**:260–268.
454 24. Ye X, Luke B, Andresson T, Blonder J: **18O Stable Isotope Labeling in MS-based**
455 **Proteomics**. *Brief Funct Genomics* 2009, **8**:136–144.
456 25. Hsu J-L, Chen S-H: **Stable isotope dimethyl labelling for quantitative proteomics**
457 **and beyond**. *Philos Trans R Soc A Math Phys Eng Sci* 2016, **374**:20150364.
458 26. Evans C, Noirel J, Ow SY, Salim M, Pereira-Medrano AG, Couto N, Pandhal J, Smith D,
459 Pham TK, Karunakaran E, et al.: **An insight into iTRAQ: where do we stand now?**
460 *Anal Bioanal Chem* 2012, **404**:1011–1027.
461 27. Bąchor R, Waliczek M, Stefanowicz P, Szewczuk Z: **Trends in the Design of New**
462 **Isobaric Labeling Reagents for Quantitative Proteomics**. *Mol* 2019, **24**.
463 28. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP: **Absolute quantification of**
464 **proteins and phosphoproteins from cell lysates by tandem MS**. *Proc Natl Acad Sci*
465 2003, **100**:6940 LP – 6945.
466 29. Kirkpatrick DS, Gerber SA, Gygi SP: **The absolute quantification strategy: a general**
467 **procedure for the quantification of proteins and post-translational modifications.**
468 *Methods* 2005, **35**:265–273.
469 30. Kettenbach AN, Rush J, Gerber SA: **Absolute quantification of protein and post-**
470 **translational modification abundance with stable isotope–labeled synthetic**
471 **peptides**. *Nat Protoc* 2011, **6**:175–186.
472 31. Picard G, Lebert D, Louwagie M, Adrait A, Huillet C, Vandenesch F, Bruley C, Garin J,
473 Jaquinod M, Brun V: **PSAQ^{TM} standards for accurate MS–based quantification of**
474 **proteins: from the concept to biomedical applications**. *J Mass Spectrom* 2012,
475 **47**:1353–1363.
476 32. Pratt JM, Simpson DM, Doherty MK, Rivers J, Gaskell SJ, Beynon RJ: **Multiplexed**
477 **absolute quantification for proteomics using concatenated signature peptides**
478 **encoded by QconCAT genes**. *Nat Protoc* 2006, **1**:1029–1043.
479 33. Beynon RJ, Doherty MK, Pratt JM, Gaskell SJ: **Multiplexed absolute quantification in**
480 **proteomics using artificial QCAT proteins of concatenated signature peptides**. *Nat*

Methods 2005, **2**:587–589.
34. Simpson DM, Beynon RJ: **QconCATs: design and expression of concatenated protein standards for multiplexed protein quantification**. *Anal Bioanal Chem* 2012, **404**:977–989.
35. Scott KB, Turko I V, Phinney KW: **Chapter Eleven - QconCAT: Internal Standard for Protein Quantification**. In *Isotope Labeling of Biomolecules - Applications*. Edited by Kelman ZBT-M in E. Academic Press; 2016:289–303.
36. O'Rourke MB, Town SEL, Dalla P V, Bicknell F, Koh Belic N, Violi JP, Steele JR, Padula MP: **What is Normalization? The Strategies Employed in Top-Down and Bottom-Up Proteome Analysis Workflows**. *Proteomes* 2019, **7**.
37. Redding-Johanson AM, Batth TS, Chan R, Krupa R, Szmidt HL, Adams PD, Keasling JD, Soon Lee T, Mukhopadhyay A, Petzold CJ: **Targeted proteomics for metabolic pathway optimization: Application to terpene production**. *Metab Eng* 2011, **13**:194–203.
38. Singh P, Batth TS, Juminaga D, Dahl RH, Keasling JD, Adams PD, Petzold CJ: **Application of targeted proteomics to metabolically engineered Escherichia coli**. *Proteomics* 2012, **12**:1289–1299.
39. George KW, Chen A, Jain A, Batth TS, Baidoo EEK, Wang G, Adams PD, Petzold CJ, Keasling JD, Lee TS: **Correlation analysis of targeted proteins and metabolites to assess and engineer microbial isopentenol production**. *Biotechnol Bioeng* 2014, **111**:1648–1658.
40. Brunk E, George KW, Alonso-Gutierrez J, Thompson M, Baidoo E, Wang G, Petzold CJ, McCloskey D, Monk J, Yang L, et al.: **Characterizing Strain Variation in Engineered E. coli Using a Multi-Omics-Based Workflow**. *Cell Syst* 2016, **2**:335–346.
41. Alonso-Gutierrez J, Kim E-M, Batth TS, Cho N, Hu Q, Chan LJG, Petzold CJ, Hillson NJ, Adams PD, Keasling JD, et al.: **Principal component analysis of proteomics (PCAP) as a tool to direct metabolic engineering**. *Metab Eng* 2015, **28**:123–133.
42. Volke DC, Rohwer J, Fischer R, Jennewein S: **Investigation of the methylerythritol 4-phosphate pathway for microbial terpenoid production through metabolic control analysis**. *Microb Cell Fact* 2019, **18**:192.
43. Zhu F, Zhong X, Hu M, Lu L, Deng Z, Liu T: **In vitro reconstitution of mevalonate pathway and targeted engineering of farnesene overproduction in Escherichia coli**. *Biotechnol Bioeng* 2014, **111**:1396–1405.
44. Zhu F, Lu L, Fu S, Zhong X, Hu M, Deng Z, Liu T: **Targeted engineering and scale up of lycopene overproduction in Escherichia coli**. *Process Biochem* 2015, **50**:341–346.
45. Liu Q, Wu K, Cheng Y, Lu L, Xiao E, Zhang Y, Deng Z, Liu T: **Engineering an iterative polyketide pathway in Escherichia coli results in single-form alkene and alkane overproduction**. *Metab Eng* 2015, **28**:82–90.
46. Gaida SM, Liedtke A, Jentges AHW, Engels B, Jennewein S: **Metabolic engineering of Clostridium cellulolyticum for the production of n-butanol from crystalline cellulose**. *Microb Cell Fact* 2016, **15**:6.
47. Bartasun P, Prandi N, Storch M, Aknin Y, Bennett M, Palma A, Baldwin G, Sakuragi Y, Jones PR, Rowland J: **The effect of modulating the quantity of enzymes in a model ethanol pathway on metabolic flux in Synechocystis sp. PCC 6803**. *PeerJ* 2019, **7**:e7529.
48. Nishiguchi H, Hiasa N, Uebayashi K, Liao J, Shimizu H, Matsuda F: **Transomics data-driven, ensemble kinetic modeling for system-level understanding and engineering of the cyanobacteria central metabolism**. *Metab Eng* 2019, **52**:273–283.
49. Yunus IS, Jones PR: **Photosynthesis-dependent biosynthesis of medium chain-length fatty acids and alcohols**. *Metab Eng* 2018, **49**:59–68.

532 50. Yunus IS, Palma A, Trudeau DL, Tawfik DS, Jones PR: **Methanol-free biosynthesis of**
533 **fatty acid methyl ester (FAME) in Synechocystis sp. PCC 6803**. *Metab Eng* 2020,
534 **57**:217–227.
535 51. Tan G-Y, Deng K, Liu X, Tao H, Chang Y, Chen J, Chen K, Sheng Z, Deng Z, Liu T:
536 **Heterologous Biosynthesis of Spinosad: An Omics-Guided Large Polyketide**
537 **Synthase Gene Cluster Reconstitution in Streptomyces**. *ACS Synth Biol* 2017,
538 **6**:995–1005.
539 52. Geiselman GM, Zhuang X, Kirby J, Tran-Gyamfi MB, Prahl J-P, Sundstrom ER, Gao Y,
540 Munoz Munoz N, Nicora CD, Clay DM, et al.: **Production of ent-kaurene from**
541 **lignocellulosic hydrolysate in Rhodosporidium toruloides**. *Microb Cell Fact* 2020,
542 **19**:24.
543 53. Petersen A, Crocoll C, Halkier BA: **De novo production of benzyl glucosinolate in**
544 **Escherichia coli**. *Metab Eng* 2019, **54**:24–34.
545 54. Janke C, Gaida S, Jennewein S: **The production of isoprene from cellulose using**
546 **recombinant Clostridium cellulolyticum strains expressing isoprene synthase**.
547 *Microbiologyopen* 2020, **9**:e1008.
548 55. Sasaki Y, Eng T, Herbert RA, Trinh J, Chen Y, Rodriguez A, Gladden J, Simmons BA,
549 Petzold CJ, Mukhopadhyay A: **Engineering Corynebacterium glutamicum to produce**
550 **the biogasoline isopentenol from plant biomass hydrolysates**. *Biotechnol Biofuels*
551 2019, **12**:41.
552 56. Hammel A, Zimmer D, Sommer F, Mühlhaus T, Schroda M: **Absolute Quantification of**
553 **Major Photosynthetic Protein Complexes in Chlamydomonas reinhardtii Using**
554 **Quantification Concatamers (QconCATs)** . *Front Plant Sci* 2018, **9**:1265.
555 57. Wienkoop S, Weiß J, May P, Kempa S, Irgang S, Recuenco-Munoz L, Pietzke M,
556 Schwemmer T, Rupprecht J, Egelhofer V, et al.: **Targeted proteomics for**
557 **Chlamydomonas reinhardtii combined with rapid subcellular protein fractionation,**
558 **metabolomics and metabolic flux analyses**. *Mol Biosyst* 2010, **6**:1018–1031.
559 58. Voges R, Corsten S, Wiechert W, Noack S: **Absolute quantification of**
560 **Corynebacterium glutamicum glycolytic and anaplerotic enzymes by QconCAT**. *J*
561 *Proteomics* 2015, **113**:366–377.
562 59. Matsuda F, Ogura T, Tomita A, Hirano I, Shimizu H: **Nano-scale liquid**
563 **chromatography coupled to tandem mass spectrometry using the multiple reaction**
564 **monitoring mode based quantitative platform for analyzing multiple enzymes**
565 **associated with central metabolic pathways of Saccharomyces cerevisiae using**
566 **ultra** . *J Biosci Bioeng* 2015, **119**:117–120.
567 60. Tokumaru Y, Uebayashi K, Toyoshima M, Osanai T, Matsuda F, Shimizu H:
568 **Comparative Targeted Proteomics of the Central Metabolism and Photosystems in**
569 **SigE Mutant Strains of Synechocystis sp. PCC 6803**. *Mol* 2018, **23**.
570 61. Vuorijoki L, Isojärvi J, Kallio P, Kouvonen P, Aro E-M, Corthals GL, Jones PR, Muth-
571 Pawlak D: **Development of a Quantitative SRM-Based Proteomics Method to Study**
572 **Iron Metabolism of Synechocystis sp. PCC 6803**. *J Proteome Res* 2016, **15**:266–279.
573 62. Toyoshima M, Sakata M, Ueno Y, Toya Y, Matsuda F, Akimoto S, Shimizu H: **Proteome**
574 **analysis of response to different spectral light irradiation in Synechocystis sp. PCC**
575 **6803**. *J Proteomics* 2021, **246**:104306.
576 63. Carroll KM, Simpson DM, Eyers CE, Knight CG, Brownridge P, Dunn WB, Winder CL,
577 Lanthaler K, Pir P, Malys N, et al.: **Absolute Quantification of the Glycolytic Pathway**
578 **in Yeast:: DEPLOYMENT OF A COMPLETE QconCAT APPROACH *** . *Mol Cell*
579 *Proteomics* 2011, **10**.
580 64. Zulak KG, Lippert DN, Kuzyk MA, Domanski D, Chou T, Borchers CH, Bohlmann J:
581 **Targeted proteomics using selected reaction monitoring reveals the induction of**
582 **specific terpene synthases in a multi-level study of methyl jasmonate-treated**

583 **Norway spruce (Picea abies)**. *Plant J* 2009, **60**:1015–1030.
584 65. Wienkoop S, Larrainzar E, Glinski M, González EM, Arrese-Igor C, Weckwerth W:
585 **Absolute quantification of Medicago truncatula sucrose synthase isoforms and N-**
586 **metabolism enzymes in symbiotic root nodules and the detection of novel nodule**
587 **phosphoproteins by mass spectrometry**. *J Exp Bot* 2008, **59**:3307–3315.
588 66. Song J, Du L, Li L, Palmer LC, Forney CF, Fillmore S, Zhang Z, Li X: **Targeted**
589 **quantitative proteomic investigation employing multiple reaction monitoring on**
590 **quantitative changes in proteins that regulate volatile biosynthesis of strawberry**
591 **fruit at different ripening stages**. *J Proteomics* 2015, **126**:288–295.
592 67. Batth TS, Singh P, Ramakrishnan VR, Sousa MML, Chan LJG, Tran HM, Luning EG, Pan
593 EHY, Vuu KM, Keasling JD, et al.: **A targeted proteomics toolkit for high-throughput**
594 **absolute quantification of Escherichia coli proteins**. *Metab Eng* 2014, **26**:48–56.
595 68. Kohlstedt M, Sappa PK, Meyer H, Maaß S, Zaprasis A, Hoffmann T, Becker J, Steil L,
596 Hecker M, van Dijl JM, et al.: **Adaptation of Bacillus subtilis carbon core metabolism**
597 **to simultaneous nutrient limitation and osmotic challenge: a multi-omics**
598 **perspective**. *Environ Microbiol* 2014, **16**:1898–1917.
599 69. Voges R, Noack S: **Quantification of proteome dynamics in Corynebacterium**
600 **glutamicum by 15N-labeling and selected reaction monitoring**. *J Proteomics* 2012,
601 **75**:2660–2669.
602 70. Schiffmann C, Hansen R, Baumann S, Kublik A, Nielsen PH, Adrian L, von Bergen M,
603 Jehmlich N, Seifert J: **Comparison of targeted peptide quantification assays for**
604 **reductive dehalogenases by selective reaction monitoring (SRM) and precursor**
605 **reaction monitoring (PRM).** *Anal Bioanal Chem* 2014, **406**:283–291.
606 71. Kito K, Okada M, Ishibashi Y, Okada S, Ito T: **A strategy for absolute proteome**
607 **quantification with mass spectrometry by hierarchical use of peptide-concatenated**
608 **standards**. *Proteomics* 2016, **16**:1457–1473.
609 72. Song J, Du L, Li L, Kalt W, Palmer LC, Fillmore S, Zhang Y, Zhang Z, Li X: **Quantitative**
610 **changes in proteins responsible for flavonoid and anthocyanin biosynthesis in**
611 **strawberry fruit at different ripening stages: A targeted quantitative proteomic**
612 **investigation employing multiple reaction monitoring**. *J Proteomics* 2015, **122**:1–10.
613 73. Gao Y, Fillmore TL, Munoz N, Bentley GJ, Johnson CW, Kim J, Meadows JA, Zucker JD,
614 Burnet MC, Lipton AK, et al.: **High-Throughput Large-Scale Targeted Proteomics**
615 **Assays for Quantifying Pathway Proteins in Pseudomonas putida KT2440** . *Front*
616 *Bioeng Biotechnol* 2020, **8**:1383.
617 74. Smallbone K, Messiha HL, Carroll KM, Winder CL, Malys N, Dunn WB, Murabito E,
618 Swainston N, Dada JO, Khan F, et al.: **A model of yeast glycolysis based on a**
619 **consistent kinetic characterisation of all its enzymes**. *FEBS Lett* 2013, **587**:2832–
620 2841.
621 75. Weaver LJ, Sousa MML, Wang G, Baidoo E, Petzold CJ, Keasling JD: **A kinetic-based**
622 **approach to understanding heterologous mevalonate pathway function in E. coli**.
623 *Biotechnol Bioeng* 2015, **112**:111–119.
624 76. Trauchessec M, Jaquinod M, Bonvalot A, Brun V, Bruley C, Ropers D, de Jong H, Garin
625 J, Bestel-Corre G, Ferro M: **Mass Spectrometry-based Workflow for Accurate**
626 **Quantification of Escherichia coli Enzymes: How Proteomics Can Play**
627 **a Key Role in Metabolic Engineering *** . *Mol Cell Proteomics* 2014, **13**:954–968.
628 77. Tao H, Zhang Y, Cao X, Deng Z, Liu T: **Absolute quantification of proteins in the fatty**
629 **acid biosynthetic pathway using protein standard absolute quantification**. *Synth*
630 *Syst Biotechnol* 2016, **1**:150–157.
631 78. Darmawi J, K. BEE, M. R-JA, S. BT, Helcio B, Aindrila M, J. PC, D. KJ: **Modular**
632 **Engineering of l-Tyrosine Production in Escherichia coli** . *Appl Environ Microbiol*
633 2012, **78**:89–98.

634 79. Philipps G, de Vries S, Jennewein S: **Development of a metabolic pathway transfer**
635 **and genomic integration system for the syngas-fermenting bacterium Clostridium**
636 **ljungdahlii**. *Biotechnol Biofuels* 2019, **12**:112.
637 80. George KW, Thompson MG, Kang A, Baidoo E, Wang G, Chan LJG, Adams PD, Petzold
638 CJ, Keasling JD, Soon Lee T: **Metabolic engineering for the high-yield production of**
639 **isoprenoid-based C5 alcohols in E. coli**. *Sci Rep* 2015, **5**:11128.
640 81. Nowroozi FF, Baidoo EEK, Ermakov S, Redding-Johanson AM, Batth TS, Petzold CJ,
641 Keasling JD: **Metabolic pathway optimization using ribosome binding site variants**
642 **and combinatorial gene assembly**. *Appl Microbiol Biotechnol* 2014, **98**:1567–1581.
643 82. Ma SM, Garcia DE, Redding-Johanson AM, Friedland GD, Chan R, Batth TS, Haliburton
644 JR, Chivian D, Keasling JD, Petzold CJ, et al.: **Optimization of a heterologous**
645 **mevalonate pathway through the use of variant HMG-CoA reductases**. *Metab Eng*
646 2011, **13**:588–597.
647 83. Zhang F, Ouellet M, Batth TS, Adams PD, Petzold CJ, Mukhopadhyay A, Keasling JD:
648 **Enhancing fatty acid production by the expression of the regulatory transcription**
649 **factor FadR**. *Metab Eng* 2012, **14**:653–660.
650 84. Wang G, Jia W, Chen N, Zhang K, Wang L, Lv P, He R, Wang M, Zhang D: **A GFP-**
651 **fusion coupling FACS platform for advancing the metabolic engineering of**
652 **filamentous fungi**. *Biotechnol Biofuels* 2018, **11**:232.
653 85. Garcia DC, Dinglasan JLN, Shrestha H, Abraham PE, Hettich RL, Doktycz MJ: **A lysate**
654 **proteome engineering strategy for enhancing cell-free metabolite production**.
655 *Metab Eng Commun* 2021, **12**:e00162.
656 86. Sánchez BJ, Zhang C, Nilsson A, Lahtvee P-J, Kerkhoven EJ, Nielsen J: **Improving the**
657 **phenotype predictions of a yeast genome-scale metabolic model by incorporating**
658 **enzymatic constraints**. *Mol Syst Biol* 2017, **13**:935.
659 87. Fang X, Lloyd CJ, Palsson BO: **Reconstructing organisms in silico: genome-scale**
660 **models and their emerging applications**. *Nat Rev Microbiol* 2020, **18**:731–743.
661 88. Liu JK, Lloyd C, Al-Bassam MM, Ebrahim A, Kim J-N, Olson C, Aksenov A, Dorrestein P,
662 Zengler K: **Predicting proteome allocation, overflow metabolism, and metal**
663 **requirements in a model acetogen**. *PLOS Comput Biol* 2019, **15**:e1006848.
664 89. Goelzer A, Muntel J, Chubukov V, Jules M, Prestel E, Nölker R, Mariadassou M,
665 Aymerich S, Hecker M, Noirot P, et al.: **Quantitative prediction of genome-wide**
666 **resource allocation in bacteria**. *Metab Eng* 2015, **32**:232–243.
667 90. Maaβ S, Wachlin G, Bernhardt J, Eymann C, Fromion V, Riedel K, Becher D, Hecker M:
668 **Highly Precise Quantification of Protein Molecules per Cell During Stress and**
669 **Starvation Responses in Bacillus subtilis***. *Mol Cell Proteomics* 2014,
670 **13**:2260–2276.
671 91. Nikel PI, de Lorenzo V: **Pseudomonas putida as a functional chassis for industrial**
672 **biocatalysis: From native biochemistry to trans-metabolism**. *Metab Eng* 2018,
673 **50**:142–155.
674 92. Kukurugya MA, Mendonca CM, Solhtalab M, Wilkes RA, Thannhauser TW, Aristilde L:
675 **Multi-omics analysis unravels a segregated metabolic flux network that tunes co-**
676 **utilization of sugar and aromatic carbons in Pseudomonas putida**. *J Biol*
677 *Chem* 2019, **294**:8464–8479.
678 93. Zhu D, Adebisi WA, Ahmad F, Sethupathy S, Danso B, Sun J: **Recent Development of**
679 **Extremophilic Bacteria and Their Application in Biorefinery**. *Front Bioeng*
680 *Biotechnol* 2020, **8**:483.
681 94. Crosby JR, Laemthong T, Lewis AM, Straub CT, Adams MWW, Kelly RM: **Extreme**
682 **thermophiles as emerging metabolic engineering platforms**. *Curr Opin Biotechnol*
683 2019, **59**:55–64.
684 95. Gregson BH, Metodieva G, Metodiev M V, Golyshin PN, McKew BA: **Protein expression**

685 **in the obligate hydrocarbon-degrading psychrophile Oleispira antarctica RB-8**
686 **during alkane degradation and cold tolerance**. *Environ Microbiol* 2020, **22**:1870–1883.
687 96. Van PT, Schmid AK, King NL, Kaur A, Pan M, Whitehead K, Koide T, Facciotti MT, Goo
688 YA, Deutsch EW, et al.: **Halobacterium salinarum NRC-1 PeptideAtlas: Toward**
689 **Strategies for Targeted Proteomics and Improved Proteome Coverage**. *J Proteome*
690 *Res* 2008, **7**:3755–3764.
691 97. Burg D, Ng C, Ting L, Cavicchioli R: **Proteomics of extremophiles**. *Environ Microbiol*
692 2011, **13**:1934–1955.
693 98. Lawson CE, Martí JM, Radivojevic T, Jonnalagadda SVR, Gentz R, Hillson NJ, Peisert S,
694 Kim J, Simmons BA, Petzold CJ, et al.: **Machine learning for metabolic engineering: A**
695 **review**. *Metab Eng* 2021, **63**:34–60.
696 99. Zhang J, Petersen SD, Radivojevic T, Ramirez A, Pérez-Manríquez A, Abeliuk E,
697 Sánchez BJ, Costello Z, Chen Y, Fero MJ, et al.: **Combining mechanistic and machine**
698 **learning models for predictive engineering and optimization of tryptophan**
699 **metabolism**. *Nat Commun* 2020, **11**:4880.
700 100. Jervis AJ, Carbonell P, Vinaixa M, Dunstan MS, Hollywood KA, Robinson CJ, Rattray
701 NJW, Yan C, Swainston N, Currin A, et al.: **Machine Learning of Designed**
702 **Translational Control Allows Predictive Pathway Optimization in Escherichia coli**.
703 *ACS Synth Biol* 2019, **8**:127–136.
704 101. Radivojević T, Costello Z, Workman K, Garcia Martin H: **A machine learning**
705 **Automated Recommendation Tool for synthetic biology**. *Nat Commun* 2020,
706 **11**:4879.
707