
Article:
Landels, A., Evans, C., Noirel, J. et al. (1 more author) (2015) Advances in proteomics for production strain analysis. Current Opinion in Biotechnology, 35. 111 - 117. ISSN 0958-1669

https://doi.org/10.1016/j.copbio.2015.05.001

Advances in proteomics for production strain analysis

Andrew Landels[a†], Caroline Evans[a†], Josselin Noirel[b] and Phillip C Wright[a].

† These authors contributed equally to the manuscript

Addresses:

[a] Department of Chemical and Biological Engineering, University of Sheffield, S1 3JD, UK.

[b] Chaire de bioinformatique, Laboratory 'genomics, bioinformatics and applications',

Conservatoire national des arts et métiers, 75003 Paris, France

Corresponding author: Phillip C Wright (p.c.wright@sheffield.ac.uk)

Highlights

- **Proteomics is widely used in production strain analysis**
- **The value of specific strategies is discussed with reference to case studies**
- **Methodologies often based on prior application to eukaryotic systems**
- **New developments target quantitative accuracy and proteome coverage**

Abstract

Proteomics is the large-scale study and analysis of proteins, directed to analysing protein function in a cellular context. Since the vast majority of the processes occurring in a living cell rely on protein activity, proteomics offer a unique vantage point from which researchers can dissect, characterise, understand and manipulate biological systems. When developing a production strain, proteomics offers a versatile toolkit of analytical techniques. In this commentary, we highlight a number of recent developments in this field using three industrially relevant case studies: targeted proteomic analysis of heterologous pathways in *E. coli*, biofuel production in *Synechocystis* PCC6803 and proteomic investigations of lignocellulose degradation. We conclude by discussing future developments in proteomics that will impact upon metabolic engineering and process monitoring of bio-producer strains.

Introduction

Chemical biotechnology is a field directed to harnessing living organisms as cellular factories, for bio-based production of small molecules and polymers [1,2]. These biological production systems are less well understood than traditional chemical engineering processes due to their inherent complexity. As a result, advanced molecular techniques like proteomics are required to engineer more efficient processes and develop new applications.

First defined in 1995 as a portmanteau of 'protein' and 'genomics', proteomics is the large-scale study of proteins within a cell, tissue or organism [3]. It is a rapidly evolving field focused on identification and characterisation of these proteins and their proteoforms (isoforms and post translational modification (PTM) variants). Quantitative methods in proteomics have enabled comparative analysis of protein expression profiles, typically providing 'snapshots' of cells and proteins in different stages of bio-production. Recent studies have also measured protein turnover by determining rates of protein synthesis and degradation. These techniques offer a means to gain information on mechanisms of bio-production for purposes of optimisation and process monitoring. To date proteomics has found application to well characterised strains such as *E. coli* [4], emergent bio-producer strains like the cyanobacteria *Synechocystis* PCC6803 (herein referred to as *Synechocystis*) [5•]; as well as metaproteomic analysis of mixed microbial communities [6].

Bibliometric analysis (see supplementary material) of recent proteomics publications has highlighted a couple of key trends in producer strain studies: Proteomics in producer strain analysis tends to focus much more on understanding mechanisms and responses, or suggesting molecular pathways, indicating that in general production analysis is lagging behind the general trend toward targeted proteomics (Fig 1.). We cover the topic of

targeted proteomics in more detail below and highlight a small number of cutting edge studies in our first case study.

In this commentary, we present a typical approach for conducting a proteomics experiment, highlighting key terms and concepts. We then outline novel proteomics approaches using post-2012 examples, focusing on three industrially relevant case studies: a method-specific approach, a strain-specific approach and a process-specific approach, concluding with a discussion of the impact of recent developments in the field.

<Figure 1>

Proteomic analysis pipeline

High-throughput proteomic methods commonly used in biotechnology approaches utilise the 'shotgun' or bottom-up technique [7•], where the proteome is digested into peptides that are typically 5 – 14 amino acids long. Whole proteins or larger polypeptides can also be analysed (top-down, middle down respectively), but this strategy has several technical issues detailed elsewhere [8••]. A digested proteome is complex, containing thousands of peptides with varying abundances. The mix requires fractionation, typically using offline high performance liquid chromatography (HPLC) or in solution isoelectric focusing, which splits the single sample into lower complexity fractions. Doing this collects together peptides with similar features – such as hydrophobicity, charge state or isoelectric point – and significantly improves the quality of the final data. Samples are then subject to nano-flow reverse phase HPLC, coupled directly to mass spectrometer. This process is referred to as MS-MS or MS^2.

The mass spectrometer (MS) initially scans the masses and intensities of all eluting peptides from the HPLC, on a scale of seconds to milliseconds, this is an MS 'survey scan'. Eluting peptides are then selected for fragmentation from the survey scan, either in a data dependent (DDA) or data independent (DIA) acquisition mode. DDA targets a specific peak from the survey scan for further analysis, whilst DIA fragments all ions from the survey scan simultaneously. The data is then analysed computationally to identify and characterise the proteome. Two detailed reviews provide further information, Altelaar et al [9••] provide an overview, whilst Zhang et al [8••] cover the topic more comprehensively.

Approaches in proteomics

Proteomic approaches can be subdivided into discovery and targeted modes, for characterisation of the proteome and analysis of an identified subset of proteins respectively. Their application and relevance are outlined in the case studies. Examples of workflows, gel and non-gel based, together with their major benefits and drawbacks and examples of their application to bio-producer strains are outlined (Table 1). Classical proteomics employs two-dimensional electrophoresis (2DE) for analysis of expression profile, where protein identification requires MS as a second step. Gel free quantification methods achieve protein identification and relative quantification using MS with significant advantages over 2DE [10]**.** Protein and peptide labeling methods, metabolic (eg SILAC) or chemical labels (eg iTRAQ) have been widely employed in proteomics but 'label-free' methods are increasingly gaining in popularity [11,12]. To date, no direct comparison of all these techniques has been reported (Table 1). Technique selection is dependent on the biological context, number of samples to be processed and compared.

Targeted proteomic approaches are directed to the detection and the precise quantification of specific subset of proteins of interest. This complements discovery proteomics and applications include verification of candidate proteins and process monitoring [13]. Quantification is based on detection and measurement of proteotypic peptides that represent the protein, based on unique amino acid sequence. Specificity and sensitivity are both conferred via 'reaction monitoring' for the presence of (co-eluting) fragment ions, linking precursor and product transition information. Application of high-resolution mass measurement and acquisition of full fragment ion spectra have enabled developments,

including higher throughput and specificity conferred by parallel reaction monitoring (PRM) as recently reviewed [14].

Inclusion of stable isotope forms of reference proteotypic peptides, at known concentrations, enables absolute quantification. QconCATs (concatenated proteotypic peptide sequences), are custom designed recombinant proteins, which can be metabolically labelled, purified and tryptically digested, to provide a set of standards for absolute quantification of multiple proteins in parallel [15•]. Label free approaches are popular due to limited sample pre-processing requirements prior to analysis compared to label based methodologies. Examples include Intensity-Based Absolute Quantification (iBAQ) and Absolute Protein Expression (APEX), which have been compared for different sample types and MS platforms [16-18].

<Table 1>

Case Study: Targeted proteomics for process optimisation

A key area of proteomic application is assessment and modelling of heterologous pathways. Whilst assessing how an inserted pathway is affecting the proteomic background provides useful information on how the organism is responding; for pathway engineering purposes it is often more informative to assess either the pathway proteins directly, or a specific subset of the proteome known to interact with it. Targeted proteomic methods like selective reaction monitoring mass spectrometry (SRM-MS) are useful for collecting highly repeatable, high-accuracy, quantitative data. These techniques are gaining popularity in bio-production pathway modelling and optimisation, as well as providing a means of assessment of standard parts and devices in synthetic biology. Proof of concept of this approach has been demonstrated for heterologous pathway expression in E coli as model/paradigm for optimization of heterologous pathways.

SRM-MS has been validated against analysis of red fluorescent protein expression levels in an expression plasmid and output of the tyrosine production pathway, controlled with a variety of different strength constitutive promoters [60]. These methods can also be coupled with quantification methods that incorporate a standard, such as QconCAT, to generate absolute protein quantification levels [15•]. A major advantage of this is that absolute protein values can be incorporated into kinetic metabolic models alongside metabolite data; whilst relative quantification values – which are more commonly associated with global proteome assessment – cannot [61]. In practical production terms, it has also been used to optimise production of biofuels such as isopentanol [62] and biosynthesis precursors like terpenes [63].

Case study: *Synechocystis* PCC6803

Cyanobacteria are a phylum of photosynthetic bacteria that offer promise in solar-powered bio-production. *Synechocystis* is a fully-sequenced, naturally transformable strain of cyanobacterium that is gaining popularity as a model production chassis [23]. It is currently being investigated for a variety of different products including precursors such as isoprenoids and lactic acid from CO_2 [24,25], as well as biofuels like ethanol, butanol and hydrogen [5•,26,27].

When engineering a strain for production, proteomic methods integrated with transcriptome data are used in about 20% of studies. These studies are frequently used to assess cellular response to production stresses (supplementary bibliometric analysis). Proteins and pathways found to be regulated in response to stresses are good candidates for forward-engineering strategies; either through the more traditional method of managing metabolic flux in the organism, or by understanding and controlling responses. An initial proteomics analysis of butanol stress in *Synechocystis* highlighted multiple simultaneous pathways activating in response to the stress [28], was followed up by a transcriptome study [29]. Joint analysis of these data identified slr1037 as part of a butanol-specific paired signal transduction system [5•]. This was verified with a knock-out, which was more robust to butanol stress whilst maintaining wild-type growth rate under standard conditions.

PTMs are a highly conserved method of regulation in biological systems. Despite the enrichment strategies required for identifying these low abundance features, it is possible to assess PTMs on a systems level [9••]. Three pioneering studies have been conducted in the last two years, cataloguing system-wide PTM responses in *Synechocystis* and

demonstrating their role in regulation of photosynthesis and central metabolic pathways [19,30,31].

Despite the advantages of proteomics, it requires tuning to the organism being studied. In *Synechocystis*, photosynthetic antennae proteins make up 20% of the proteome by mass [32]. This results in a large dynamic range relative to other producer-strain bacteria, such as *E. coli*, and so the advantage conferred by light harvesting capacity comes with the negative limit to proteomic coverage. This problem exists in any case where a small number of proteins are present at a very high abundance, relative to the rest of the sample such as the case with RuBisCO in plants [33]. Work has been carried out to reduce the abundance of these antenna proteins for better production and improved proteomic coverage [59]. The dynamic range problem can be alleviated to an extent through depletion strategies and use of high resolution, high throughput MS.

Case study: lignocellulose degradation

Lignocellulose is a complex polysaccharide constituent of plant cell wall. It is a promising substrate for production of molecules like ethanol and lactic acid; however, lignin inhibits the action of many common enzymes by sequestering the cellulose and xylose. This creates a bottleneck in efficient bio-production from this material [34]. Proteomics is at the forefront of deciphering solutions to this problem, either through fully integrated systems analysis or assessment of solutions utilising multiple organisms simultaneously.

Integration of metabolite, transcript and protein data, termed systems analysis, can be used to generate comprehensive models for how a cell is responding. The protein data offers an impression of the current state of the cell, whilst the transcript analysis identifies responses at high-coverage coverage and the metabolite data give an impression of flux [35,54]. Proteomics data from several lignocellulose degradation investigations, including *Clostridium* and a variety of filamentous fungi, have been integrated with transcript and metabolite data to understand how activated pathways affect cellular dynamics [36,37]. This approach has also been used to assess how xylose as a carbon source affects metabolism in yeast, to design better production strategies [38•].

The xylose catabolism study utilises pre-existing models for assembled pathways, and is typical of a proteomic work-flow for pathway analysis where identified and quantified proteins are overlaid on a metabolic model. Models are assembled using pathways found with literature analysis and new models are typically constructed by modifying an existing model from a similar organism. A detailed commentary of advancements in this process is provided by King et al in this issue [39]. Where a pre-existing model is lacking, other general exploratory assessments, like principal component analysis, can be used instead to look for

general trends in protein expression data [40]. This technique can highlight proteins, or entire conditions, that cluster together depending on the focus of the study; with suitable experimental design this can provide information in lieu of a completed metabolic model. Due to the complexity of even seemingly simple strains, data interpretation still is a limiting factor in systems level analyses. This has led to the use of cutting-edge informatics, such as machine learning, being employed in analysis of the data [41].

Whilst individual organisms have shown effectiveness in degrading lignocellulose, natural systems utilise a combination organisms performing distinct roles to achieve the effect more efficiently [42]. The study of these communities through the analysis of the proteins is referred to as 'metaproteomics', where the community is profiled instead of focusing on specific organisms, or specific pathways [43]. A growing number of studies are being carried out using metaproteomics towards the ultimate aim of engineering these systems through 'Synthetic Ecology' [44]. This emerging field will likely be of importance to the bio-production community in the future.

Addressing the challenges and perspectives

Proteomics is a very dynamic area: the proteomic toolkit is constantly expanding with both the development of both novel technologies alongside new uses of existing technologies. In the latter category, development of quantitative proteomic technologies with higher multiplexing capability, neutron encoded TMT and SILAC reagents, improves the multiplexing capability of TMT [45,46]. This enables both higher throughput and the additional benefits conferred by increased sample replication. Metabolic labels such as SILAC can be directed not only to expression profiling, but also to protein dynamics [47]. The Super-SILAC approach, involves mixing samples from different conditions to generate an internal standard for cross sample comparison, and can be combined with iBAQ to give absolute copy-number level protein quantitation [48••]**.**

Improved sample preparation strategies improve protein identification rates and coverage. In general, reagents that are very effective at solubilising proteins are incompatible with MS. The technique of Filter Aided Sample Preparation (FASP), allows removal of detergents such as SDS and other contaminants. The FASP technique, first described in 2009 [49] has been adapted for improved proteolytic digestion (eFASP) [50] and for compatibility with chemical labelling strategies for analysis of proteins (iFASP) [51] or affinity purified protein complexes (abFASP) [52]. Protein-protein interactions (PPI) is a growth area in proteomic analysis, particularly since proteins typically function in complexes which are temporally and spatially dynamic within the cell [53] and analysis of PPI has value in system based network modelling [54]. Novel strategies include enhanced capability of protein-protein interaction analysis by use of affinity enrichment MS, as an alternative to classical ‘pull down’ approaches [55].

In terms of MS instrumentation, performance is continually improving in terms of speed, sensitivity and resolution, such that complete proteome coverage is achievable [56]. Unlike DDA, DIA methods are inherently not applicable to use of metabolic or chemical labels. This limitation is being overcome with approaches such as NeuCoDIA for multiplex analysis [57]. In general, whilst MS instrumentation was directed to operating in discovery or targeted modes, next generation instruments offer both capabilities. For example, DIA methods are enabling discovery and targeted modes of data analysis to be integrated, by retrospective 'MRM like' mining of discovery data for specific peptides. With MS developments there is a need for corresponding capabilities in processing software to fully mine the increasingly complex proteomic datasets.

Conclusions

In this commentary, we have highlighted using bibliometric analysis of the field and specific case studies that proteomics is widely used in production strain analysis. Due to the dynamic nature of the field, there are a number of cutting-edge developments that have improved quantitative proteomics over the review period and are continuing to emerge. These have generated step-changes in technical and conceptual approaches to process optimisation, in both single species as well as microbial communities [6,58].

Acknowledgements

A.L and P.C.W are funded by the European Union Seventh Program for research, technical development and demonstration for funding under grant agreement No 308518, CyanoFactory. C.E. and P.C.W. thank the EPSRC for funding (EP/E036252/1). J.N. is funded by the Fondation Recherche Médicale (FRM: ING20140129444).

References and recommended reading

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

- • Of special interest
- •• Of outstanding interest

1. Fischer CR, Schaffer S: **Editorial overview: chemical biotechnology: the expansion of chemical biotechnology**. *Curr Opin Biotechnol* 2014, **30**:v-vii.
2. Becker J, Wittmann C: **Advanced biotechnology: metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products**. *Angew Chem Int Ed Engl* 2015, **54**:3328-3350.
3. Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humphery-Smith I: **Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium**. *Electrophoresis* 1995, **16**:1090-1094.
4. Wisniewski JR, Rakus D: **Multi-enzyme digestion FASP and the 'Total Protein Approach'-based absolute quantification of the Escherichia coli proteome**. *J Proteomics* 2014, **109**:322-331.
5. • Chen L, Wu L, Wang J, Zhang W: **Butanol tolerance regulated by a two-component response regulator Slr1037 in photosynthetic Synechocystis sp. PCC 6803**. In *Biotechnol Biofuels*; 2014:89.

This study provides a clear example of a follow-up investigation to a proteomics-led discovery, following identification of targets using a combination of proteomics and transcriptomics under butanol stress.

6. Abraham PE, Giannone RJ, Xiong W, Hettich RL: **Metaproteomics: extracting and mining proteome information to characterize metabolic activities in microbial communities**. *Curr Protoc Bioinformatics* 2014, **46**:13 26 11-13 26 14.
7. • Yates JR, 3rd: **The revolution and evolution of shotgun proteomics for large-scale proteome analysis**. *J Am Chem Soc* 2013, **135**:1629-1640.

This review provides a background on the progression of shotgun proteomics. This is an interesting account of the historical progression of the field and nicely frames the subject.

8. •• Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR, 3rd: **Protein analysis by shotgun/bottom-up proteomics**. *Chem Rev* 2013, **113**:2343-2394.

This comprehensive and detailed review covers a wide range of proteomics-related topics and discusses depths beyond the scope of this commentary. We highly recommend this review to readers who would like to learn more about a variety of topics in proteomics.

9. •• Altelaar AF, Munoz J, Heck AJ: **Next-generation proteomics: towards an integrative view of proteome dynamics**. In *Nat Rev Genet*; 2013:35-48.
This review provides an excellent introduction to the topic. We highly recommend it as background reading for readers who are interested by proteomics but are unfamiliar with the subject matter being discussed in this commentary.

10. Gygi SP, Corthals GL, Zhang Y, Rochon Y, Aebersold R: **Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology**. In *Proc Natl Acad Sci U S A*; 2000:9390-9395.
11. Evans C, Noirel J, Ow SY, Salim M, Pereira-Medrano AG, Couto N, Pandhal J, Smith D, Pham TK, Karunakaran E, et al.: **An insight into iTRAQ: where do we stand now?** *Anal Bioanal Chem* 2012, **404**:1011-1027.
12. Christoforou AL, Lilley KS: **Isobaric tagging approaches in quantitative proteomics: the ups and downs**. *Anal Bioanal Chem* 2012, **404**:1029-1037.
13. George KW, Alonso-Gutierrez J, Keasling JD, Lee TS: **Isoprenoid Drugs, Biofuels, and Chemicals-Artemisinin, Farnesene, and Beyond**. *Adv Biochem Eng Biotechnol* 2015.
14. Lesur A, Ancheva L, Kim YJ, Berchem G, van Oostrum J, Domon B: **Screening protein isoforms predictive for cancer using immuno-affinity capture and fast LC-MS in PRM mode**. *Proteomics Clin Appl* 2015.
15. • Batth TS, Singh P, Ramakrishnan VR, Sousa MM, Chan LJ, Tran HM, Luning EG, Pan EH, Vuu KM, Keasling JD, et al.: **A targeted proteomics toolkit for high-throughput absolute quantification of Escherichia coli proteins**. *Metab Eng* 2014, **26C**:48-56.
This study highlights the creation of a tool for high-speed, high-accuracy quantification of proteins in *E. coli*. It demonstrates how synthetic biology approaches to proteomics are championing high-accuracy quantification data at the expense of producing a large number of identifications.

16. Arike L, Valgepea K, Peil L, Nahku R, Adamberg K, Vilu R: **Comparison and applications of label-free absolute proteome quantification methods on Escherichia coli**. In *J Proteomics*. 2012 Elsevier B.V; 2012:5437-5448.
17. Ahrne E, Molzahn L, Glatter T, Schmidt A: **Critical assessment of proteome-wide label-free absolute abundance estimation strategies**. *Proteomics* 2013, **13**:2567-2578.
18. Krey JF, Wilmarth PA, Shin JB, Klimek J, Sherman NE, Jeffery ED, Choi D, David LL, Barr-Gillespie PG: **Accurate label-free protein quantitation with high- and low-resolution mass spectrometers**. *J Proteome Res* 2014, **13**:1034-1044.
19. Mikkat S, Fulda S, Hagemann M: **A 2D gel electrophoresis-based snapshot of the phosphoproteome in the cyanobacterium Synechocystis sp. strain PCC 6803**. In *Microbiology*; 2014:296-306.
20. Ciesielska K, Li B, Groeneboer S, Van Bogaert I, Lin YC, Soetaert W, Van de Peer Y, Devreese B: **SILAC-based proteome analysis of Starmerella bombicola sophorolipid production**. *J Proteome Res* 2013, **12**:4376-4392.
21. Tang X, Feng H, Zhang J, Chen WN: **Comparative proteomics analysis of engineered Saccharomyces cerevisiae with enhanced biofuel precursor production**. In *PLoS One*; 2013:e84661.
22. Yap TW, Rabu A, Abu Bakar FD, Rahim RA, Mahadi NM, Illias RM, Murad AM: **Growth phase-dependent proteomes of the Malaysian isolated Lactococcus lactis dairy strain M4 using**

label-free qualitative shotgun proteomics analysis. *ScientificWorldJournal* 2014, **2014**:642891.
23. Machado IM, Atsumi S: **Cyanobacterial biofuel production**. *J Biotechnol* 2012, **162**:50-56.
24. Englund E, Pattanaik B, Ubhayasekera SJ, Stensjo K, Bergquist J, Lindberg P: **Production of squalene in Synechocystis sp. PCC 6803**. In *PLoS One*; 2014:e90270.
25. Angermayr SA, van der Woude AD, Correddu D, Vreugdenhil A, Verrone V, Hellingwerf KJ: **Exploring metabolic engineering design principles for the photosynthetic production of lactic acid by Synechocystis sp. PCC6803**. In *Biotechnol Biofuels*; 2014:99.
26. Qiao J, Wang J, Chen L, Tian X, Huang S, Ren X, Zhang W: **Quantitative iTRAQ LC-MS/MS proteomics reveals metabolic responses to biofuel ethanol in cyanobacterial Synechocystis sp. PCC 6803**. *J Proteome Res* 2012, **11**:5286-5300.
27. Pinto F, van Elburg KA, Pacheco CC, Lopo M, Noirel J, Montagud A, Urchueguia JF, Wright PC, Tamagnini P: **Construction of a chassis for hydrogen production: physiological and molecular characterization of a Synechocystis sp. PCC 6803 mutant lacking a functional bidirectional hydrogenase**. In *Microbiology*; 2012:448-464.
28. Tian X, Chen L, Wang J, Qiao J, Zhang W: **Quantitative proteomics reveals dynamic responses of Synechocystis sp. PCC 6803 to next-generation biofuel butanol**. *J Proteomics* 2013, **78**:326-345.
29. Zhu H, Ren X, Wang J, Song Z, Shi M, Qiao J, Tian X, Liu J, Chen L, Zhang W: **Integrated OMICS guided engineering of biofuel butanol-tolerance in photosynthetic Synechocystis sp. PCC 6803**. In *Biotechnol Biofuels*; 2013:106.
30. Chardonnet S, Sakr S, Cassier-Chauvat C, Le Marechal P, Chauvat F, Lemaire SD, Decottignies P: **First proteomic study of s-glutathionylation in cyanobacteria**. *J Proteome Res* 2015, **14**:59-71.
31. Mo R, Yang M, Chen Z, Cheng Z, Yi X, Li C, He C, Xiong Q, Chen H, Wang Q, et al.: **Acetylome Analysis Reveals the Involvement of Lysine Acetylation in Photosynthesis and Carbon Metabolism in the Model Cyanobacterium Synechocystis sp. PCC 6803**. *J Proteome Res* 2015, **14**:1275-1286.
32. Gan CS, Reardon KF, Wright PC: **Comparison of protein and peptide prefractionation methods for the shotgun proteomic analysis of Synechocystis sp. PCC 6803**. *Proteomics* 2005, **5**:2468-2478.
33. Gupta R, Wang Y, Agrawal GK, Rakwal R, Jo IH, Bang KH, Kim ST: **Time to dig deep into the plant proteome: a hunt for low-abundance proteins**. *Front Plant Sci* 2015, **6**:22.
34. Zeng Y, Zhao S, Yang S, Ding SY: **Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels**. *Curr Opin Biotechnol* 2014, **27**:38-45.
35. Huang L, Kim D, Liu X, Myers CR, Locasale JW: **Estimating relative changes of metabolic fluxes**. In *PLoS Comput Biol*; 2014:e1003958.
36. Schellenberg JJ, Verbeke TJ, McQueen P, Krokhin OV, Zhang X, Alvare G, Fristensky B, Thallinger GG, Henrissat B, Wilkins JA, et al.: **Enhanced whole genome sequence and annotation of Clostridium stercorarium DSM8532T using RNA-seq transcriptomics and high-throughput proteomics**. In *BMC Genomics*; 2014:567.
37. Klaubauf S, Narang HM, Post H, Zhou M, Brunner K, Mach-Aigner AR, Mach RL, Heck AJ, Altelaar AF, de Vries RP: **Similar is not the same: differences in the function of the (hemi-)cellulolytic regulator XlnR (Xlr1/Xyr1) in filamentous fungi**. *Fungal Genet Biol* 2014, **72**:73-81.
38. • Latimer LN, Lee ME, Medina-Cleghorn D, Kohnz RA, Nomura DK, Dueber JE: **Employing a combinatorial expression approach to characterize xylose utilization in Saccharomyces cerevisiae**. *Metab Eng* 2014, **25**:20-29.

This study uses proteomics and metabolomics to determine how engineered S. cerevisiae catabolises xylose – a sugar created in the breakdown of lignocellulose. It neatly frames how

proteomics can be used to determine changes in pathway flux to determine the fate of metabolites in an engineered system.

39. King ZA, Lloyd CJ, Feist AM, Palsson Bo: **Next-generation genome-scale models for metabolic engineering**. *Current Opinion in Biotechnology* 2015, **35**:23–29.
40. Alonso-Gutierrez J, Kim EM, Batth TS, Cho N, Hu Q, Chan LJ, 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.
41. Kelchtermans P, Bittremieux W, De Grave K, Degroeve S, Ramon J, Laukens K, Valkenborg D, Barsnes H, Martens L: **Machine learning applications in proteomics research: how the past can boost the future**. *Proteomics* 2014, **14**:353-366.
42. Boaro AA, Kim YM, Konopka AE, Callister SJ, Ahring BK: **Integrated 'omics analysis for studying the microbial community response to a pH perturbation of a cellulose-degrading bioreactor culture**. *FEMS Microbiol Ecol* 2014, **90**:802-815.
43. Vanwonterghem I, Jensen PD, Ho DP, Batstone DJ, Tyson GW: **Linking microbial community structure, interactions and function in anaerobic digesters using new molecular techniques**. *Curr Opin Biotechnol* 2014, **27**:55-64.
44. Pandhal J, Noirel J: **Synthetic microbial ecosystems for biotechnology**. *Biotechnol Lett* 2014, **36**:1141-1151.
45. Werner T, Sweetman G, Savitski MF, Mathieson T, Bantscheff M, Savitski MM: **Ion coalescence of neutron encoded TMT 10-plex reporter ions**. *Anal Chem* 2014, **86**:3594-3601.
46. Merrill AE, Hebert AS, MacGilvray ME, Rose CM, Bailey DJ, Bradley JC, Wood WW, El Masri M, Westphall MS, Gasch AP, et al.: **NeuCode labels for relative protein quantification**. In *Mol Cell Proteomics*. 2014 by The American Society for Biochemistry and Molecular Biology, Inc.; 2014:2503-2512.
47. Trotschel C, Albaum SP, Wolff D, Schroder S, Goesmann A, Nattkemper TW, Poetsch A: **Protein turnover quantification in a multilabeling approach: from data calculation to evaluation**. In *Mol Cell Proteomics*; 2012:512-526.
48.•• Soufi B, Krug K, Harst A, Macek B: **Characterization of the E. coli proteome and its modifications during growth and ethanol stress**. *Front Microbiol* 2015, **6**:103.

This demonstrates the use of super-SILAC, a cutting edge quantitative proteomics strategy in *E. coli* under ethanol biofuel production stress.

49. Wisniewski JR, Zougman A, Nagaraj N, Mann M: **Universal sample preparation method for proteome analysis**. In *Nat Methods*; 2009:359-362.
50. Nel AJ, Garnett S, Blackburn JM, Soares NC: **Comparative Reevaluation of FASP and Enhanced FASP Methods by LC-MS/MS**. *J Proteome Res* 2015, **14**:1637-1642.
51. McDowell GS, Gaun A, Steen H: **iFASP: combining isobaric mass tagging with filter-aided sample preparation**. *J Proteome Res* 2013, **12**:3809-3812.
52. Huber ML, Sacco R, Parapatics K, Skucha A, Khamina K, Muller AC, Rudashevskaya EL, Bennett KL: **abFASP-MS: affinity-based filter-aided sample preparation mass spectrometry for quantitative analysis of chemically labeled protein complexes**. *J Proteome Res* 2014, **13**:1147-1155.
53. Wright PC, Jaffe S, Noirel J, Zou X: **Opportunities for protein interaction network-guided cellular engineering**. *IUBMB Life* 2013, **65**:17-27.
54. Kasavi C, Eraslan S, Arga KY, Oner ET, Kirdar B: **A system based network approach to ethanol tolerance in Saccharomyces cerevisiae**. In *BMC Syst Biol*; 2014:90.
55. Keilhauer EC, Hein MY, Mann M: **Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS)**. In *Mol Cell Proteomics*. 2015 by The American Society for Biochemistry and Molecular Biology, Inc.; 2015:120-135.

56. Mann M, Kulak NA, Nagaraj N, Cox J: **The coming age of complete, accurate, and ubiquitous proteomes**. *Mol Cell* 2013, **49**:583-590.
57. Minogue CE, Hebert AS, Rensvold JW, Westphall MS, Pagliarini DJ, Coon JJ: **Multiplexed quantification for data-independent acquisition**. *Anal Chem* 2015, **87**:2570-2575.
58. Pan C, Banfield JF: **Quantitative metaproteomics: functional insights into microbial communities**. *Methods Mol Biol* 2014, **1096**:231-240.
59. Kwon J-H, Bernat G, Wagner H, Roegner M, Rexroth S: **Reduced light-harvesting antenna: Consequences on cyanobacterial metabolism and photosynthetic productivity**. *Algal Research-Biomass Biofuels and Bioproducts* 2013, **2**:188-195.
60. 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.
61. Weaver LJ, Sousa MML, Wang G, Baidoo E, Petzold CJ, Keasling JD: **A Kinetic-Based Approach to Understanding Heterologous Mevalonate Pathway Function in E. Coli**. *Biotechnology and Bioengineering* 2015, **112**:111-119.
62. 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**. *Biotechnology and Bioengineering* 2014, **111**:1648-1658.
63. Redding-Johanson AM, Batth TS, Chan R, Krupa R, Szmidt HL, Adams PD, Keasling JD, Lee TS, Mukhopadhyay A, Petzold CJ: **Targeted proteomics for metabolic pathway optimization: Application to terpene production**. *Metabolic Engineering* 2011, **13**:194-203.
64. Li Y, Han F, Xu H, Mu J, Chen D, Feng B, Zeng H: **Potential lipid accumulation and growth characteristic of the green alga Chlorella with combination cultivation mode of nitrogen (N) and phosphorus (P)**. *Bioresource Technology* 2014, **174**:24-32.