Omics studies elucidate an increase in bio-production efficiency

High performance proteomics and metabolomics studies using the impact II Q-TOF provide insight into how rational strain design increased bio-production efficiency.

This application note describes a combination of metabolomic and proteomic studies that provided an understanding into how the efficiency of arginine production by *C. glutamicum*, a gram negative bacterium used for the biotechnological production of primary metabolites, was improved by rational strain design. The results obtained by the Bruker impact II Q-TOF system using the InstantExpertise™ data-dependent HRAM LC-MS/MS acquisition routine provided accurate and reliable identification and relative quantitation for the proteomics and metabolomics studies. Pathway mapping of the proteomics data clearly showed an increase in the abundance of enzymes involved in arginine biosynthesis in the mutant strains when compared to the wild type. The presence of bottlenecks could only be deciphered in combination with the metabolomics data, highlighting the importance of combined metabolomics and proteomics studies.

Keywords:
impact II
Pathway mapping
Multi-Omics
Proteomics
Metabolomics
MetaboScape
InstantExpertise™

Authors: Frederik Walter, Marcus Persicke, Jörn Kalinowski
Bielefeld University, Bielefeld, Germany
Stephanie Kaspar-Schoenefeld, Aiko Barsch, Heiko Neuweger, Matthias Szesny, Nikolas Kessler
Bruker Daltonik GmbH, Bremen, Germany

Introduction

Because many of today's nutritional and consumer products are produced by biotechnology, there is growing interest in the development of new techniques that improve production efficiency. Rational strain design is one approach researchers consider when they want to improve the yield of amino acids such as arginine, an important amino acid widely used by the cosmetic, food and pharmaceutical industries.

This application note describes a powerful technique for increasing the efficiency of arginine production by *C. glutamicum*, a gram negative bacterium used for the biotechnological production of primary metabolites, using rational strain design and a combination of metabolomic and proteomic studies.

Because the Bruker impact II Q-TOF system is well suited to bottom-up proteomic and metabolomic workflows, it was used to acquire the metabolomics and proteomics data. A connection between mass spectrometric data and biology was established by mapping the results of both metabolic and proteomic changes onto the biochemical pathways. The results of this study provide a deeper understanding of how rational strain design can be used to increase the efficiency of arginine production. The workflows described can also be extended to the study of other biotechnological production systems.

Results

A combination of metabolomics and proteomics studies was used to decipher the changes in arginine production caused by the introduction of rational strain design in *C. glutamicum*. Three different mutant strains were examined, all of which had an influence on the arginine biosynthesis pathway (see Figure 1).

Data was acquired on the Bruker impact II Q-TOF system utilizing "one shot data acquisition" of MS & MS/MS spectra from all precursor ions by data dependent MS/MS applying the InstantExpertise™ routine (Figure 2). Metabolomics data processing, including the automatic identification of known targets, statistical analyses, and pathway mapping of metabolite fold changes, was conducted using Bruker's MetaboScape software. This workflow-based, client-server software automatically identified several known compounds in the arginine biosynthetic pathway with confidence based on user-definable levels for mass accuracy, isotopic fidelity, and retention time. The corresponding results were displayed as an "annotation quality". A more detailed

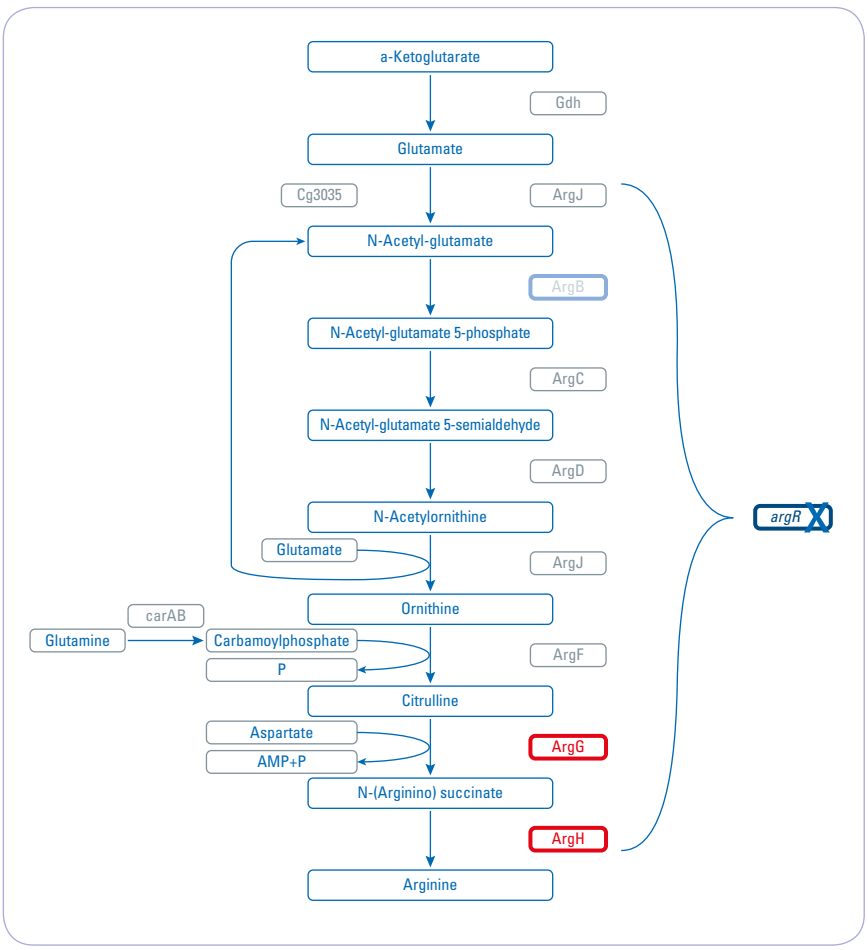

Fig. 1: Overview about arginine biosynthesis pathway and analyzed mutant strains. Three different mutant strains (argR, argB, and argGH) influencing different parts of the arginine biosynthesis pathway are investigated to increase arginine production.

- *Genes are regulated by the repressor argR, which is activated by arginine*
- ***Repressor deletion mutant (argR):** chromosomal deletion of argR.*
- ***Feedback resistant mutant (argB):** chromosomal deletion of argR and introduction of feedback-resistant $argB^{fbr}$ allels. The N-acetylglutamate kinase, which is encoded by $argB^{fb}$ is feedback regulated by arginine and thus also limits arginine production. Previous studies showed that besides arginine, these strains accumulated significant amounts of citrulline as a by-product, indicating a bottleneck in the pathway.*
- ***Debottlenecking by overexpression (argGH):** chromosomal deletion of argR and introduction of feedback-resistant $argB^{fbr}$ allels as well as debottlenecking of the last two reactions by overexpression of argGH*

description of this non-targeted High Resolution Accurate Mass (HRAM) LC-QTOF based metabolomics workflow can be found in Bruker Poster Note (PN-26 [1]).

Statistical analysis of the metabolomics data revealed a clear separation of *C. glutamicum* wild type (WT) extracts from all three mutant strains via 3D PCA score plots (Figure 3a). The "feedback resistant" (*argB*) and the "debottlenecking by overexpression" (*argGH*) mutant extracts overlapped in the PCA scores plot. Proteomics data was processed using MaxQuant software [2]. Subsequent statistical data evaluation using Perseus software [3] showed separation of the wild type extract from all three mutant strains in 3D PCA scores plot (Figure 3b); which confirmed the separation already detected in the metabolomics study. The metabolomics data showed an overlap of the "feedback resistant" and the "debottlenecking by overexpression" mutant extracts, but in the proteomics results all four of the different bacterial strains were separated.

The proteomics study revealed significant changes of proteins involved in the arginine biosynthesis pathway in the "repressor deletion mutant". In all mutants, gene deletion of *argR* should be present resulting in the absence of the corresponding protein. As expected, the data showed that the protein could be clearly identified in the wild type with eight peptides. Two peptides in the mutant lines (ELGELLSGR, AA 161 – 169 and SLGSTPSTPENLNPVTR, AA 2 – 18) were also identified. This finding agrees at the proteome level as well since only a partial deletion was genetically engineered in the *argR* gene, resulting in the mutant strain still producing amino acids 1 to 38 and 154 to 171.

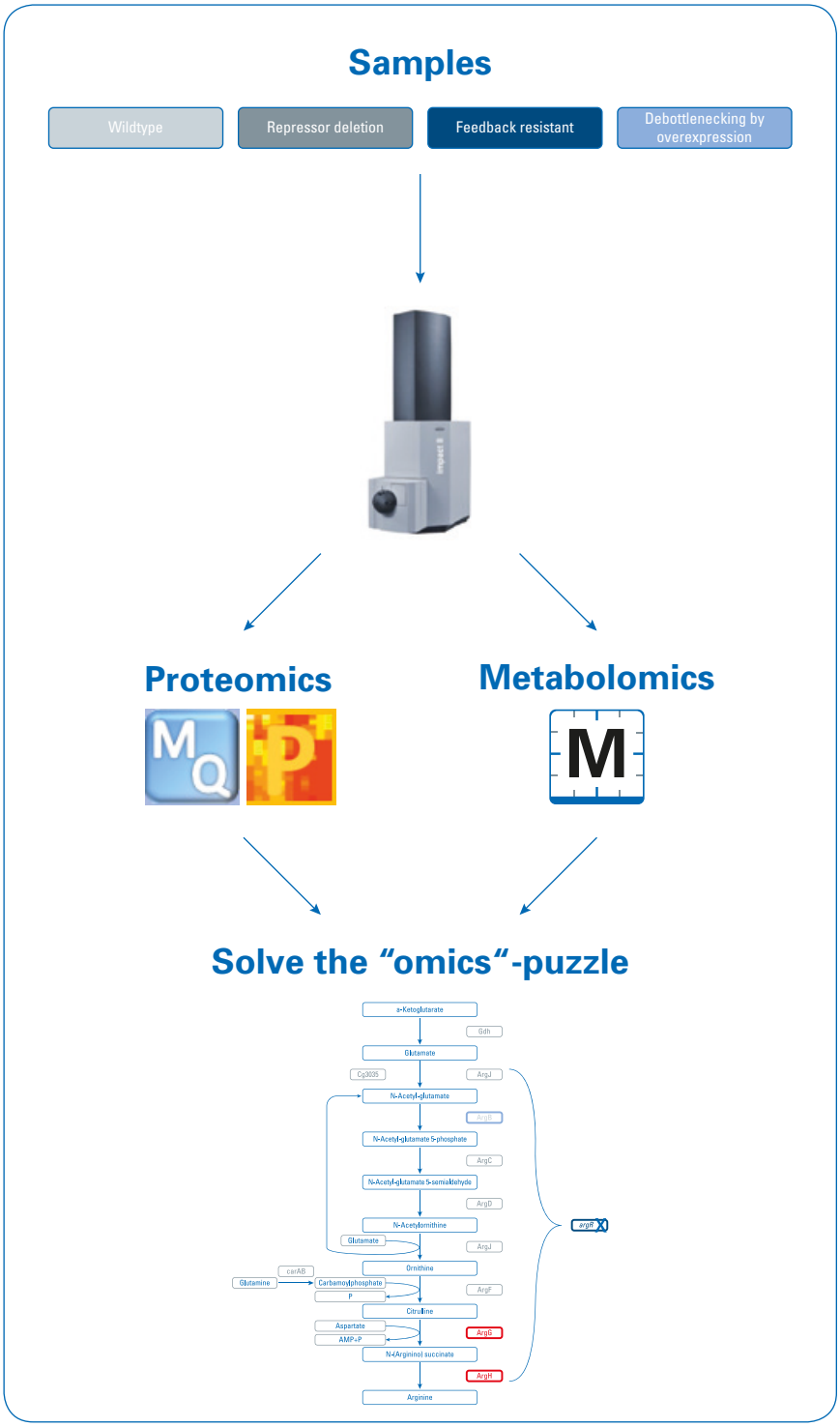

Fig. 2: Workflow to investigate the increase in arginine production by rational strain design. Four samples (wild type and three mutant strains) were analyzed using proteomics and metabolomics applications on the impact II Q-TOF instrument. Data was processed using MaxQuant and Perseus (for proteomics studies) or MetaboScape (for metabolomics studies) software then it was mapped onto the biological pathways to visualize the results of the introduced changes.

Mapping the alterations detected in the metabolomics and proteomics studies onto biochemical pathways allowed the quick formulation of a hypothesis for the observed changes in the biological context. Since the goal was to increase arginine production (Figure 4), the focus was primarily on the arginine biosynthesis pathway even though changes were observed for other pathways.

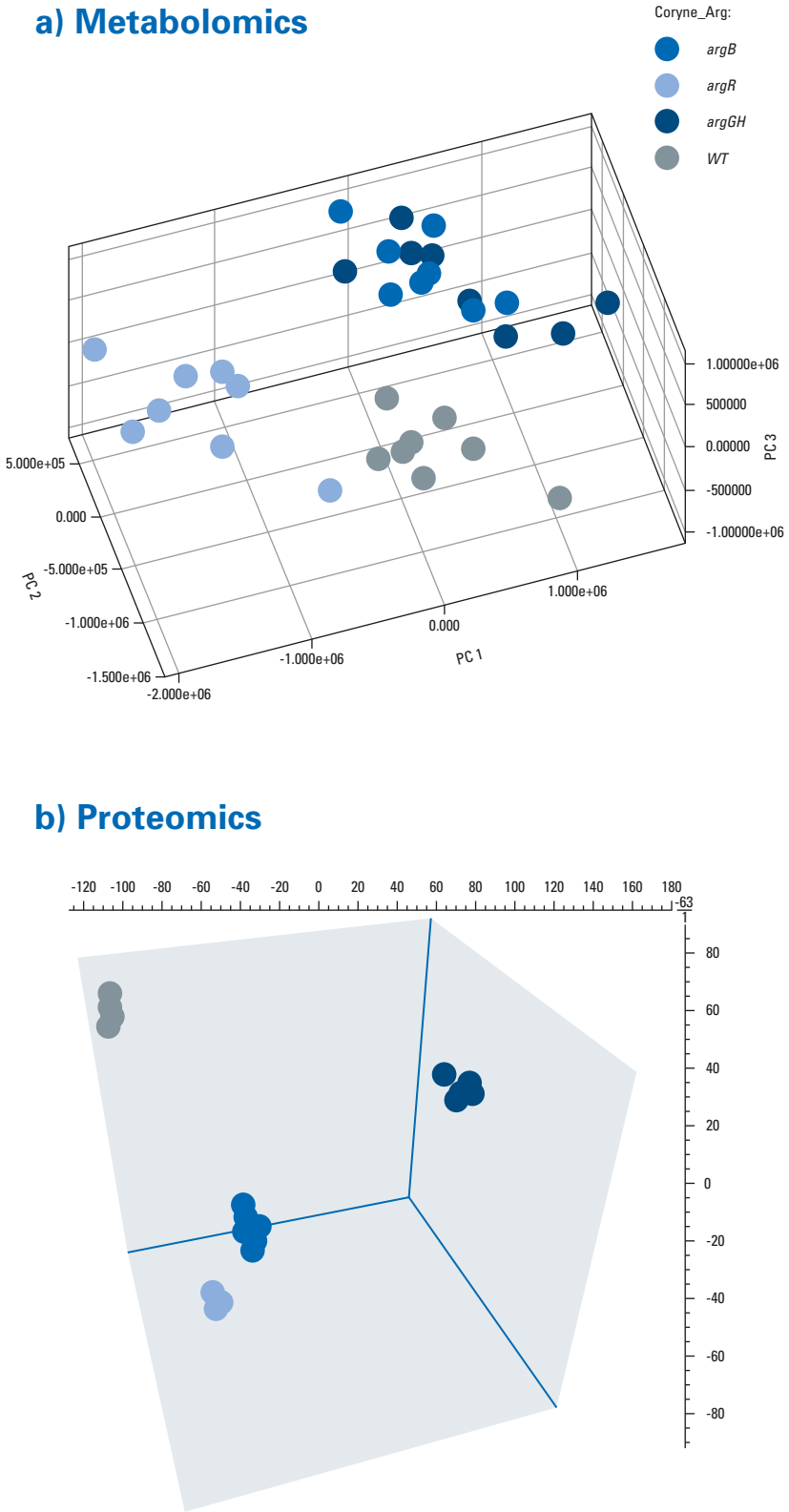

Fig. 3: PCA 3D scores for metabolomics and proteomics data processing. A clear separation between the wild type and mutant strains was detected using both approaches. The "feedback resistant" and the "debottlenecking by overexpression" mutant extracts overlap in the metabolomics PCA scores plot, and they were clearly separated in the proteomics approach.

The mapping of log2 fold changes was accomplished using custom pathway maps created in PathVisio software. The genes for biosynthesis of arginine are organized in an operon structure *(argCJBDFRGH)* and regulated by the repressor ArgR, which is activated by arginine, in the wild type. A comparison of the repressor deletion mutant with the wild type (Figure 4 a) using label-free quantitative proteomics clearly showed a significant increase in the abundance of the enzymes involved in the arginine biosynthetic pathway, with log2 fold changes greater than 3. Even though the increase in the abundance of these enzymes might be expected to induce a significant increase in arginine production, metabolomics data revealed that the depression of arginine biosynthesis inhibitors alone did not increase intracellular arginine.

As arginine was secreted into the fermentation broth, the levels of extracellular arginine levels were expected to increase. Arginine level in the fermentation broth was also analyzed (Figure 5). The extracellular levels of arginine were not detectable in the WT and remained low with levels of 38 mg/L in the repressor deletion mutant *(argR)*.

It is known that N-acetylglutamate kinase, which is encoded by *argB*, is feedback regulated by arginine. As this regulation can limit arginine production, a second mutant containing the chromosomal deletion of *argR* and an introduction of a feedback-resistant *argB* allele was created and analyzed *(argB)*. Interestingly, introduction of the feedback resistance results in reduced abundance of N-acetylglutamate kinase, encoded by the *argB* gene, in the mutant line compared to the wild type. When comparing this feedback resistant mutant to wild type (Figure 4 b) for the metabolomics

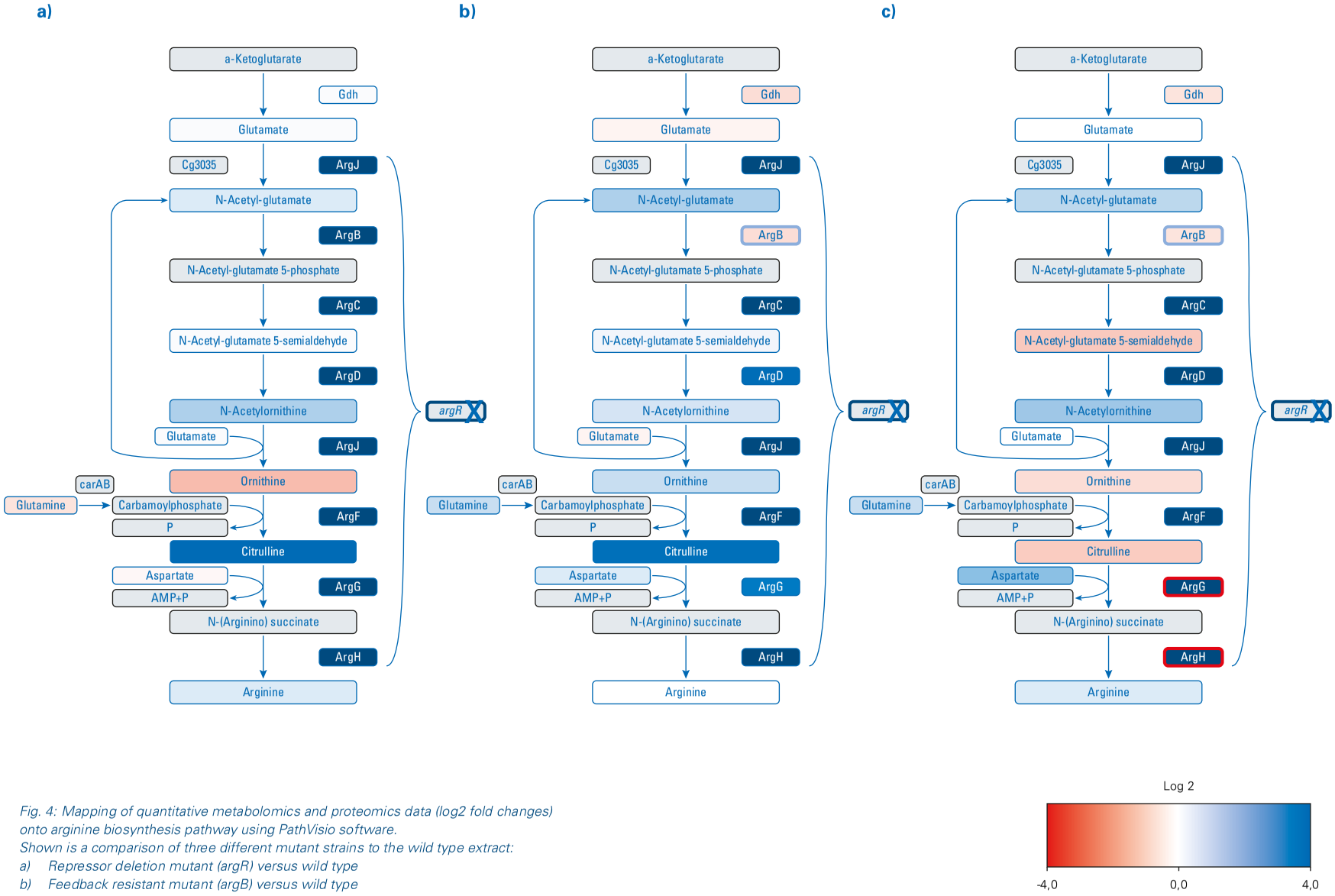

Fig. 4: Mapping of quantitative metabolomics and proteomics data (log2 fold changes) onto arginine biosynthesis pathway using PathVisio software.
Shown is a comparison of three different mutant strains to the wild type extract:

a) Repressor deletion mutant (argR) versus wild type
b) Feedback resistant mutant (argB) versus wild type
c) Debottlenecking by overexpression mutant (argGH) versus wild type

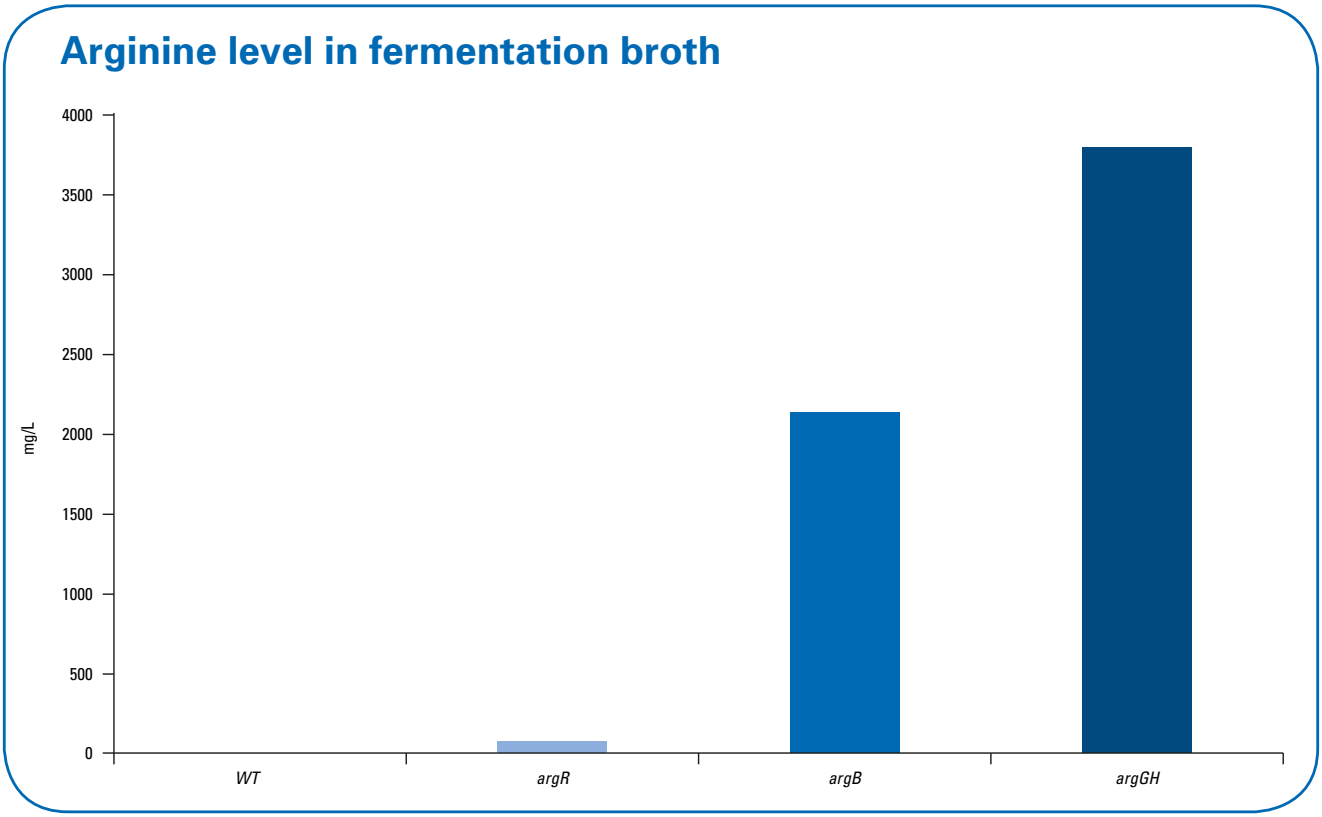

Fig. 5: 5: Extracellular levels of arginine in the fermentation broth in the wild type (wt) and the three different mutant strains. The highest levels of arginine in the fermentation broth were detected in debottlenecking by the overexpression mutant.

data, no increase in intracellular levels of arginine was observed. But, when measuring the extracellular levels of arginine, an amount of more than 2 g/L was determined – a significant increase compared to amounts determined for wild type and *argR* deletion mutant (Figure 5). Interestingly, mapping of the intracellular metabolite levels revealed an increase of the citrulline level, as highlighted by the green color coding in Figure 4b. This observation indicates a limitation in the last two steps of the pathway starting from *argG* / citrulline in the "feedback resistant mutant".

To potentially "de-bottleneck" the last two reactions of the arginine biosynthesis, a third mutant was constructed by overexpressing the *argGH* genes in the "feedback resistant mutant" background. An increase in abundance of arginosuccinate synthetase (ArgG) and arginosuccinate lyase (ArgH) was detected using label-free proteomics (Figure 4c), verifying that this overexpression actually led to an increase in protein abundance. At the same time this debottlenecking resulted in lower intracellular levels of ornithine and citrulline. Pathway mapping results also showed that intracellular levels of arginine were slightly increased. Most importantly,

Metabolomics

HPLC	Dionex™ RSLC™ (Thermo Fisher Scientific)
Column	ZIC™ HILIC 100 × 2.1 mm, 3.5 µm, PEEK, SeQuant (Merck)
Flow rate	0.3 mL/min
Mobile phase	A = water + 0.1% formic acid, B = acetonitrile + 0.1% formic acid
Gradient	t = 0–1 min, 90% B; t = 13 min 2% B; t = 16 min, 2% B; t = 17 min 90% B; t =25 min 90% B
Source	ESI source
Ionization	ESI(+)
Mode	InstantExpertise™
Scan range	m/z 20–1000
Acquisition rate	MS: 1 Hz, MSMS: 8Hz

the extracellular levels of arginine increased to more than 3.5 g/L.

Methods

Three mutant strains were compared to *C. glutamicum* wild type extracts. Metabolomics and proteomics data were acquired on an impact II Q-TOF MS (Bruker Daltonics) instrument using data-dependent MS/MS methods. MetaboScape 1.0 software was used for processing of metabolomics data, whereas MaxQuant software (v. 1.5.2.8) was used to process proteomics data. Mapping of detected changes to the arginine pathway was done in MetaboScape and PathVisio 3.2.0 [4, 5] software to enable the interpretation of data in a biological context. As arginine was secreted to the fermentation broth, extracellular arginine levels were measured using fluorescence detection [6].

Proteomics	
HPLC	Dionex™ RSLCnano™ (Thermo Fisher Scientific)
Column	Nano Trap Column, 100 µm i.d. × 2 cm, packed with Acclaim PepMap100 C_{18}, 5 µm, 100 Å, nanoViper and Acclaim PepMap RSLC C_{18}, 2 µm, 100 Å, 75 µm i.d.×50 cm, nanoViper (Thermo Fisher Scientific)
Flow rate	400 nL/min
Mobile phase	A = water + 0.1% formic acid, B = acetonitrile + 0.1% formic acid
Gradient	t = 0 min, 2% B; t = 2 min 5% B; t = 90 min, 35% B; t = 100 – 110 min 95% B; t =111 – 125 min 5% B
Source	CaptiveSpray ionization source with nanoBooster (using acetonitrile as dopant)
Ionization	ESI(+)
Mode	InstantExpertise™
Scan range	m/z 300–1600
Acquisition rate	MS: 2 Hz, MSMS: 4 – 16 Hz

Arginine measurements	
Derivatization	o-phthaldialdehyde
HPLC	Smartline Amino acid Analyser (Knauer)
Column	0.3 ACCQtaq 150 x 3.9 mm, 4 µm (Waters) mL/min
Mobile phase	A = 19% methanol in ammonium acetate buffer (pH 7.2), B = 75% methanol in ammonium acetate buffer (ph 7.2)
Gradient	t = 0 – 20 min, 0% B; t = 55 min 100% B; t = 60 min, 100% B
Detection	Excitation: 330 nm, emission: 450 nm

Sample

Three mutant strains (Δ*argR*, Δ*argR argB*^{fbr}, Δ*argR argB*^{fbr} pZ8-1::[*argGH*]) were compared to *C. glutamicum* wild type extracts.

Conclusions

- The combination of non-targeted omics techniques provided valuable insights into the changes in *C. glutamicum* production caused by rational strain design.
- The results obtained by the impact II Q-TOF system using the InstantExpertise data-dependent HRAM LC-MS/MS acquisition routine provided accurate and reliable identification and relative quantitation for the proteomics and metabolomics studies.
- Pathway mapping of the proteomics data clearly showed an increase in the abundance of enzymes involved in arginine biosynthesis in the mutant strains when compared to the wild type.
- However, the reasons for the changes in arginine production, and the presence of bottlenecks in the biosynthetic pathway, could only be deciphered in combination with the metabolomics data, highlighting the importance of combined metabolomics and proteomics

Learn More

You are looking for further information? Check out the link or scan the QR Code for our latest Webinar.

www.bruker.com/webinar-metabolomics-1116

References

[1] PN-26: MetaboScape: Linking HRAM QTOF Data to Biology – Increasing Arginine Production in *C. glutamicum* by Rational Strain Design and Discovery Metabolomics

https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/Separations_MassSpectrometry/Literature/PosterNotes/PN-26_MetaboScape_Coryne_Arginine_Mutanten_eBook.pdf

[2] Cox, J. and Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol, 2008, 26, pp 1367-72.

[3] Tyanova S., et al., Nat Methods, 2016, doi: 10.1038/nmeth.390

[4] PathVisio 3: An Extendable Pathway Analysis Toolbox
Kutmon M, van Iersel MP, Bohler A, Kelder T, Nunes N, Pico AR, and Evelo, CT
PLoS Computational Biology (2015) 11(2): e1004085
doi: 10.1371/journal.pcbi.1004085, pubmed: 25706687

[5] Presenting and exploring biological pathways with PathVisio
van Iersel MP, Kelder T, Pico AR, Hanspers K, Coort S, Conklin BR, and Evelo C.
BMC Bioinformatics. 2008 Sep 25;9:399.
doi: 10.1186/1471-2105-9-399, pubmed: 18817533

[6] Walter F., et al. ; J Biotechnol. 2015, doi: 10.1016

Bruker Daltonik GmbH

Bremen · Germany
Phone +49 (0)421-2205-0

Bruker Scientific LLC

Billerica, MA · USA
Phone +1 (978) 663-3660

ms.sales.bdal@bruker.com – www.bruker.com