Applied Microbiology and Biotechnology (2022) 106:6139–6156
https://doi.org/10.1007/s00253-022-12098-4

GENOMICS, TRANSCRIPTOMICS, PROTEOMICS

# An integrative-omics analysis of an industrial clavulanic acid-overproducing *Streptomyces clavuligerus*

Aslıhan Kurt-Kızıldoğan[1] · Gözde Çelik[2] · Eser Ünsaldı[2,3] · Servet Özcan[4] · Şerife Ayaz-Güner[5,6] · Gülay Özcengiz[2]

Received: 6 March 2022 / Revised: 21 July 2022 / Accepted: 23 July 2022 / Published online: 10 August 2022


## Abstract

Clavulanic acid (CA) is a clinically important secondary metabolite used to treat infectious diseases. We aimed to decipher complex regulatory mechanisms acting in CA biosynthesis by analyzing transcriptome- and proteome-wide alterations in an industrial CA overproducer *Streptomyces clavuligerus* strain, namely DEPA and its wild-type counterpart NRRL3585. A total of 924 differentially expressed genes (DEGs) and 271 differentially produced proteins (DPPs) were obtained by RNA-seq and nanoLC-MS/MS analyses, respectively. In particular, CA biosynthetic genes, namely, *car* (*cad*), *cas2*, *oat2*, *pah*, *bls*, *ceas2, orf12*, and *claR*, a cluster situated regulatory (CSR) gene, were significantly upregulated as shown by RNA-seq. Enzymes of clavam biosynthesis were downregulated considerably in the DEPA strain, while the genes involved in the arginine biosynthesis, one of the precursors of CA pathway, were overexpressed. However, the biosynthesis of the other CA precursor, glyceraldehyde-3-phosphate (G3P), was not affected. CA overproduction in the DEPA strain was correlated with BldD, BldG, BldM, and BldN (AdsA) overrepresentation. In addition, TetR, WhiB, and Xre family transcriptional regulators were shown to be significantly overrepresented. Several uncharacterized/unknown proteins differentially expressed in the DEPA strain await further studies for functional characterization. Correlation analysis indicated an acceptable degree of consistency between the transcriptome and proteome data. The study represents the first integrative-omics analysis in a CA overproducer *S. clavuligerus* strain, providing insights into the critical control points and potential rational engineering targets for a purposeful increase of CA yields in strain improvement.

## Key points

- *Transcriptome and proteome-wide alterations in industrial CA overproducer strain DEPA*
- *An acceptable degree of consistency between the transcriptome and proteome data*
- *New targets to be exploited for rational engineering*

**Keywords** *Streptomyces clavuligerus* DEPA · Clavulanic acid overproduction · RNA-seq · nanoLC-MS/MS

✉ Gülay Özcengiz
ozcengiz@metu.edu.tr

[1] Department of Agricultural Biotechnology, Ondokuz Mayıs University, Samsun 55139, Turkey

[2] Department of Biological Sciences, Middle East Technical University, Ankara 06800, Turkey

[3] Present Address: Gen-Era Diagnostics, Ankara 06830, Turkey

[4] Department of Biology, Erciyes University, Kayseri 38280, Turkey

[5] Department of Molecular Biology and Biotechnology, Abdullah Gul University, Kayseri 38080, Turkey

[6] Department of Molecular Biology and Genetics, Cancer Genetics and Molecular, Izmir Institute of Technology, Izmir 35430, Turkey

## Introduction

*Streptomyces clavuligerus* is a Gram-positive, spore-forming bacterium with the ability to produce a large number of secondary metabolites including the nonclassical β-lactamase inhibitor clavulanic acid (CA), clavam metabolites, the β-lactam antibiotic cephamycin C (CephC), tunicamycin-like nucleoside antibiotics, and dithiolopyrrolones, as holomycin (Nett et al. 2009). CA (3R,5R clavam) is a potent irreversible inhibitor of many class A and class D β-lactamases (Drawz and Bonomo 2010; Özcengiz and Demain 2013).

Three distinct gene clusters are defined for CA and clavam biosynthesis in *S. clavuligerus*. The CA and clavams gene clusters are located in the chromosome, while a clavams paralog gene cluster is positioned in the pSCL4

megaplasmid (Medema et al. 2010) (Figure S1). Some genes in the CA biosynthetic gene cluster are not crucial for CA biosynthesis and some late stages in CA biosynthesis are still obscure (AbuSara et al. 2019) (Figure S2).

Regulation of CA biosynthesis in *S. clavuligerus* involves a highly complex cascade in which cluster-situated regulators (CSRs) act together with global and signaling networks (Liras et al. 2008; Song et al. 2010; Paradkar 2013; Ferguson et al. 2016; Álvarez-Álvarez et al. 2017; Ünsaldı et al. 2017). CcaR, the CSR of CephC-CA supercluster, is located in the CephC cluster and positively affects CA and CephC biosynthesis. ClaR, located in the CA gene cluster, is a transcriptional activator of CA biosynthesis late steps. BldD is a master developmental repressor targeting *bldA*, *bldM*, *bldN (adsA)*, and *adpA* and exerts an indirect role in secondary metabolite biosynthesis (Bush et al. 2015). BldD plays a role in antibiotic production as an activator in two *Streptomyces* strains (Yan et al. 2020). BldG, BldA, AdpA, and Brp are global regulators controlling CA biosynthesis at different levels. *bldA* or *bldG* deletion in *S. clavuligerus* resulted in the underrepresentation of Cas2, OppA1,GcaS, and Bls2 (only in *bldG* mutant) (Ferguson et al. 2016). The AraC-type transcriptional regulator AdpA mediates the expression of both *ccaR* and early genes (*ceaS2/bls2* and *cas2*) in the CA biosynthetic gene cluster. In its absence, CA and CephC biosynthesis were drastically reduced in *S. clavuligerus* (López-García et al. 2010). The butyrolactone receptor protein, Brp, binds to ARE (autoregulatory elements) boxes in the promoter regions of *ccaR* and *adpA*, acting as a repressor. A higher level of CA production was recorded when this gene was deleted (Santamarta et al. 2005). The AreB protein plays a role in CA-CephC biosynthesis by controlling *ccaR* and *brp* expression (Santamarta et al. 2007). Besides, two-component regulatory systems and nutritional condition-based regulations affect partially CA production (Fu et al. 2019; Pinilla et al. 2019). In spite of the studies conducted, the regulatory mechanisms in CA biosynthetic pathway have not been fully elucidated (Pinilla et al. 2019).

Omics technology has contributed to understand the complex regulatory mechanisms underlying CA overproduction in *S. clavuligerus*. For instance, Medema et al. (2011) provided a brief report on genome-wide gene expression changes in a CA-overproducing industrial strain generated from ATCC 27064 type strain to show 10 upregulated operons or genes of primary and secondary metabolism as the keys to the rational design of CA overproducers. A proteomic analysis using iTRAQ of *S. clavuligerus* NRRL 3585 and its *bldG*- and *bldA*-deleted mutants was performed by Ferguson et al. (2016). Next, a gel-based comparative proteomics study was conducted by our group (Ünsaldı et al. 2017) to decipher the changes occurring in the industrial CA-overproducer *S. clavuligerus* DEPA strain developed through iterative mutagenesis. In addition, the transcriptome of *S. clavuligerus* ATCC 27064 was compared under favorable and restrictive nutritional conditions (Pinilla et al. 2019). The use of favorable complex media conditions led to the upregulated expression of many regulatory and biosynthetic CA genes. Hwang et al. (2019) presented a high-quality genome sequence of *S. clavuligerus* ATCC 27064 with 58 potential secondary metabolite biosynthetic gene clusters (BCGs) and integrating differential RNA sequencing (RNA-Seq) and ribosome profiling, they determined transcriptional and translational regulatory elements and the potential regulons of β-lactam biosynthesis.

The combination of transcriptomics, proteomics, and physiological experiments provides a comprehensive framework for systems biology (Liras and Martín 2021). In the present study, RNA-seq and nanoLC-MS/MS were combined for the first time to provide a comprehensive understanding of complex changes in mRNA and protein levels that coincide with CA overproduction in an industrial *S. clavuligerus* strain. The data presented herein will pave the way to develop novel rational engineering strategies in favor of CA overproduction in *S. clavuligerus.*

## Materials and methods

### Bacterial strains and culture conditions

*S. clavuligerus* NRRL3585 (Higgens and Kastner 1971) and its derived industrial CA overproducer strain DEPA (DEPA Pharmaceuticals Co., İzmit, Türkiye) were used in this study. Both strains were grown in Trypticase Soy Broth (TSB) (Oxoid) at 28 °C and 220 rpm in baffled flasks. TSB seed cultures were prepared as stated previously (Ünsaldı et al. 2017). Twenty-five milliliters of these seed cultures was used to inoculate 500 mL of baffled flasks containing 100 mL of semi-defined Starch-Asparagine (SA) (Aharonowitz and Demain 1979) medium for fermentation.

### Determination of growth, antibiotic yields, and residual inorganic phosphate

The growth of the cultures was determined by DNA quantification according to Burton (1968) and CA, CephC, tunicamycin, and holomycin yields by bioassay as described previously (de la Fuente et al. 2002; Özcengiz et al. 2010; Baş et al. 2020).

Residual inorganic phosphate concentration in the cultures was determined in Agilent Technologies/7700X ICP-MS systems by inductively coupled plasma – mass spectrometer (ICP-MS) analysis using culture supernatants as samples and SA medium as control.

## RNA extraction and purification

For RNA isolation, 2 mL of samples from *S. clavuligerus* NRRL3585 and DEPA were taken from 48-h old SA cultures. RNA was extracted with a Nucleospin® RNA Kit (Macherey–Nagel GmbH & Co, Germany). RNA samples were incubated with DNase I (Macherey–Nagel GmbH & Co, Germany) to avoid DNA contamination. The RNA quality and quantity were measured spectrophotometrically in a NanoDrop (BioDrop, UK). The RNA integrity was determined using the Agilent 2100 Bioanalyzer (USA) instrument. The RNAs were stabilized using RNAstable (Biomatrica, San Diego, CA, USA) before RNA-seq experiments.

## Library construction, RNA sequencing, and bioinformatic analyses

Only RNAs with RNA integrity values above 7.0 were used in the following mRNA library preparation. RNA sequencing of *S. clavuligerus* strains, NRRL3585 and DEPA, was performed on Illumina HiSeq 4000 system (BGI Genomics Co., Ltd., Hong Kong), and $2 \times 150$ bp paired-end libraries were obtained. Each sample produced an average of 3.86 Gb data. Raw RNA-seq data were deposited in the National Center for Biotechnology Information's (NCBI) Sequence Read Archive (SRA) under the accession number SRP276705. The linker sequences, low-quality reads, and residual rRNA sequences were removed from the raw data. The clean reads were mapped to *Streptomyces clavuligerus* F613-1 reference genome sequence, available in Genbank (accession numbers CP016559 and CP016560) by Bowtie2 (Langmead and Salzberg 2012) and compared using HISAT (Kim et al. 2015). IGV 2.3.32 (Thorvaldsdóttir et al. 2013) tool was used for detailed analysis of the clean data. To calculate gene expression levels, RSEM software (Li and Dewey 2011) was chosen. The expression level of each gene was normalized by the number of fragments per kilobase of transcriptome per million mapped reads (FPKM). The data were screened with the false discovery rate (FDR) and p-value thresholds of $\leq 0.05$ with a cut-off for fold change $\log_2(\mathrm{FC}) \geq 1$ to consider them as significantly differentially expressed. Differentially expressed genes (DEGs) were categorized, and pathway analysis of DEGs was performed through the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al. 2008) database. Gene Ontology (GO) terms were used for functional analysis of DEGs. Transcriptomic data were deposited in the Sequence Read Archive (SRA) database (http://www.ncbi.nlm.nih.gov/sra/) under accession number SRP276705.

## Protein extraction, filter-aided sample preparation, and Zip-Tip purification

Samples taken from 48-h SA cultures were used in protein extraction as described in Faurobert et al. (2007). The protein concentrations were determined by a modified Bradford assay according to Ramagli and Rodriguez (1985). Prior to nanoLC-MS/MS analysis, the extracted proteins were treated with protein extraction, filter-aided sample preparation (FASP™) Protein Digestion Kit (Expedeon). The resultant peptide mixtures were purified by the use of Zip-tip® Millipore™ cleaning tips (http://www.millipore.com/catalogue/module/c5737) as previously described (Ünsaldı et al. 2021).

## Nanoscale liquid chromatography coupled to tandem mass spectrometry analysis

Nanoscale liquid chromatography coupled to tandem mass spectrometry (nanoLC-MS/MS) analysis was performed as described by Ünsaldı et al. (2021). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (Deutsch et al. 2020) via the PRIDE (Perez-Riverol et al. 2022) partner repository with the dataset identifier PXD032272.

The peptide identification (Analysis of LC–MS/MS data) was carried out using the Paragon search algorithm provided by ProteinPilot 4.5 Beta (AB SCIEX) software. As the parameters of the database search, a set of biological modifications for a standard iodoacetamide cys alkylation was selected, and 0.05 Da for-precursor ion mass and 0.1 Da for-fragment mass tolerance were set. False discovery routine of the identification software was applied. Minimum of two distinct peptides and 95% unused peptide score were set as a threshold for quantification of each protein. The number of product ions detected for each peptide's identification was selected as a measure of the spectral count, and quantification calculations were applied using a modified formula described in Byrum et al. (2013). The modified version of the normalized spectral abundance factor (NSAF) for each identified protein is calculated by dividing each protein's ratio of the total number of fragment ions produced and its sequence length by the same ratio summed for all identified proteins. For the proteomics data, the proteins with a cut-off for fold change $\log_2(\mathrm{FC}) \geq 0.5$ were considered as significantly differently produced (DPP). Quantification analysis and correlation plots were produced using the R (ver. 4.0.2) statistical programming environment.

## Databases for protein categorization, function, and association for nanoLC-MS/MS

*S. clavuligerus* strain F613-1 reference proteome sequence data file was obtained from NCBI's genome

assembly repository with the accession GCA_001693675.1 ASM169367v1. UniProt database (http://www.uniprot.org/) was used for extracting *S. clavuligerus* protein data. Functional categories of the proteins were determined using the COG database (ftp://ftp.ncbi.nih.gov/pub/COG/COG2014/static/lists/listStrbin.html) while the cellular localizations were determined by PSORTb version 3.0.2 (http://www.psort.org/psortb/), Gpos-mPloc (http://www.csbio.sjtu.edu.cn/bioinf/Gpos-multi/#) and ngLoc (http://genome.unmc.edu/ngLOC/index.html). STRING 10.0 (http://STRING-db.org/) and KEGG mapper (http://www.kegg.jp/kegg/tool/annotate_sequence.html) tools were used to predict the association of proteins and pathways, respectively. Theoretical pI and Mw values of the proteins were calculated by Expasy pI/Mw tool (Gasteiger et al. 2003).

### Real-time-quantitative polymerase chain reaction

Primers used in real-time-quantitative polymerase chain reaction (RT-qPCR) experiments were designed using the OligoAnalyzer Tool (idtdna.com) and tabulated in Table S1. All experiments and statistical analyses performed in this section were given in detail in our previous report (Ünsaldı et al. 2021).

### Protein and RNA correlation analysis

Transcriptome and proteome expression data sets were compared using the correlation test with the nonparametric Spearman ranking method with a minimum 95% confidence level. The built-in function provided by the R base package is used to execute the tests and the results were plotted using the R ggplots2 package. The DPPs and corresponding transcription products were also compared by the same methodology.

## Results

### General characterization of the RNA-seq data

#### RNA sequencing, differential expression analysis, and clustering of DEGs

A total of 40.73 Mb of raw data was obtained for *S. clavuligerus* strains, NRRL3585 and DEPA. After filtering the reads, 38.73 Mb of clean data was obtained for both strains. Each sample produced an average of 3.86 Gb data, with an average ratio of 57.05% compared with the reference genome. According to gene expression analysis, 6115 genes were co-expressed in both strains, while the number of genes expressed only in DEPA and NRRL3585 strains was found to be 60 and 188, respectively (Fig. 1a). Differential gene expression (DEGs) analysis showed that 924 genes were identified as DEGs in *S. clavuligerus* DEPA compared to the strain NRRL3585. Of these, 372 were significantly upregulated, and 552 were downregulated (Fig. 1 b, c).

#### KEGG pathway enrichment analysis of DEGs

We next performed the KEGG pathway enrichment analysis for the DEGs to determine which biological pathways were significantly different in *S. clavuligerus* DEPA. Antibiotic biosynthesis, ABC transporters, amino acid metabolism, pentose phosphate pathway, arginine biosynthesis, CA biosynthesis, and oxocarboxylic acid metabolism were among the top 20 most enriched pathways (Fig. 2). Regarding KEGG pathway classification, a total of 415 DEGs were determined. From them, 124 DEGs were involved in global and overview maps in the metabolism category (Fig. 3a). The carbon metabolism category included a high number of DEGs (Fig. 3b).

#### GO functional classification of DEGs

The GO classification includes biological process, cellular component, and molecular function. To sum up, the highest DEG numbers were present in catalytic activity (318), binding (220), metabolic (166) and cellular (137) processes, membrane (168), and membrane part (158) classifications (Fig. 4 a, b)..

### General characterization of the nanoLC-MS/MS data

nanoLC-MS/MS analysis identified 464 proteins for *S. clavuligerus* DEPA and 1187 proteins for *S. clavuligerus* NRRL 3585 with at least two unique peptides (Fig. 5a). Among 441 common proteins, 154 proteins had at least 41.42% higher quantitation score for DEPA strain compared with the wild type and 117 proteins' quantitation scores were lower than the other groups with the same rate, and both groups of proteins were the candidates of differentially produced proteins (DPPs) (Fig. 5b). In view of the relative quantitative distribution of DPPs to the functional categories in DEPA strain, "Translation, ribosomal structure, and biogenesis" (21.4%), Secondary metabolite biosynthesis", and "Transport and catabolism" (15.6%) were the main categories of the over-represented ones (Fig. 5c). As to the underrepresented proteins in this strain, the most relevant categories were "Amino acid transport and metabolism" (20.5%), "Energy production and conversion" (12.8%), and "Translation, ribosomal structure and biogenesis" (9.4%) (Fig. 5d). The distribution of DPPs into the functional categories is illustrated in Fig. 5e.

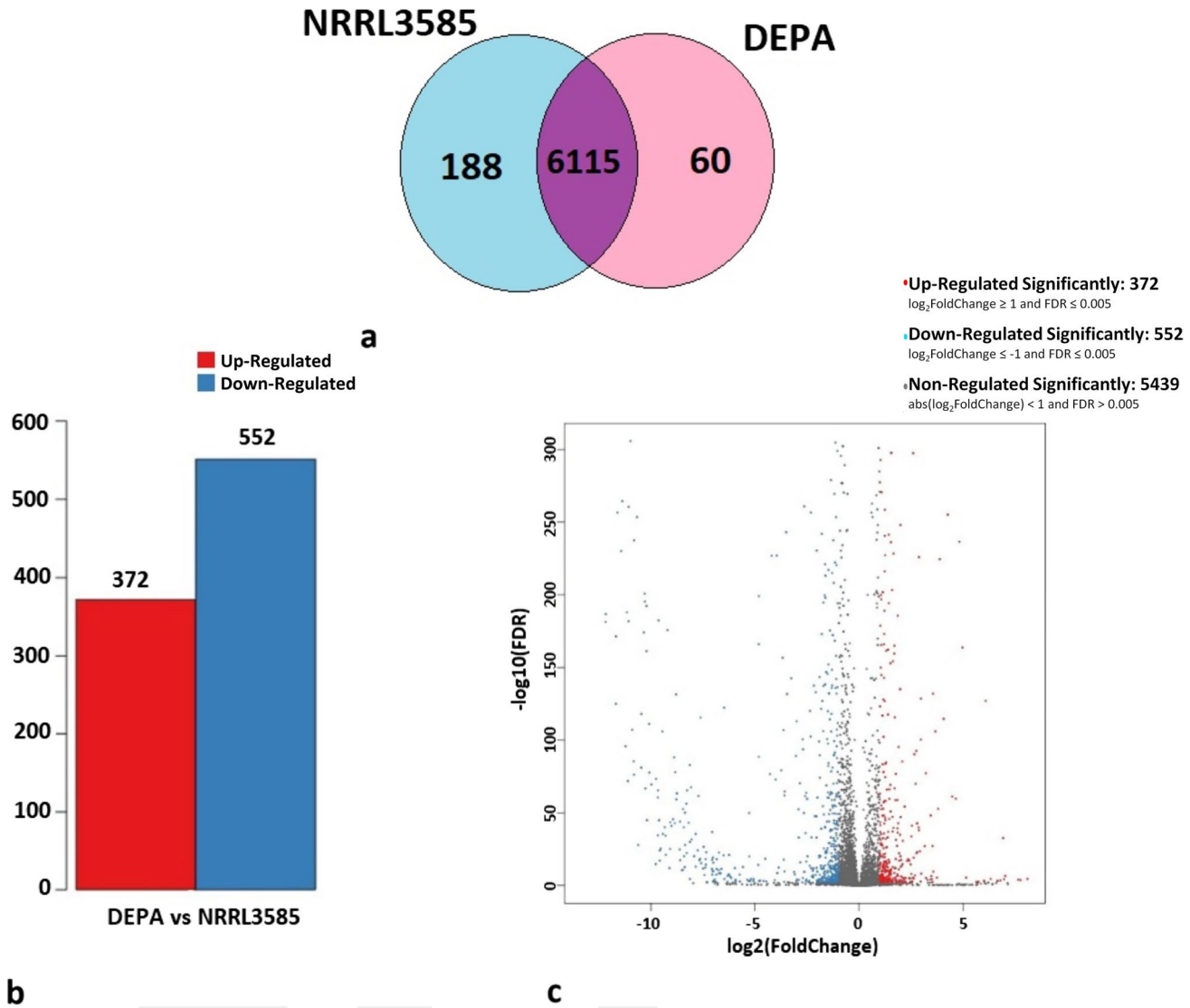


**Fig. 1** A Venn diagram showing common gene expression between *S. clavuligerus* NRRL3585 and strain DEPA (**a**). A bar graph showing statistics of DEGs in *S. clavuligerus* DEPA (**b**), and a volcano plot (**c**) showing differentially expressed genes with statistical significance in DEPA strain. The *X*-axis represents the value of $\log_2$(FC) value, and the *Y*-axis represents the value of $-\log_{10}$(FDR)

## Identification of the most significant DEGs in the DEPA strain

### The most highly expressed genes in the DEPA strain

The most highly expressed genes in the CA overproducer DEPA strain are tabulated in Table 1, excluding those coding for several hypothetical proteins. *SCLAV_1861* encoding CRISPR-associated endoribonuclease is one of the most upregulated DEGs, with 7.7-fold increase in its expression. The second most upregulated gene (6.3-fold) is *SCLAV_5662* (*rmlC*), encoding an enzyme of the dTDP-L-Rha pathway. Also, highly upregulated transcriptional regulators are belonging to TetR, TetR/AcrR family, and the developmental activator WhiB. These regulators are encoded by *SCLAV_p0892* (TetR; 5.4-fold), *SCLAV_5443* and *SCLAV_p0891* (TetR/AcrR family; 4.9- and 3.6-fold), and *SCLAV_2475 (whiB*; 5.7-fold). Concomitantly, many ABC-type transporter system genes including *SCLAV_4954* (*amfB*) (3.7-fold) that encodes a membrane translocator were highly upregulated. The related *amfT* gene (*SCLAV_4953*) coding for a serine/threonine-protein kinase was also 4.3-fold upregulated. Additionally, *bldM* and *adsA* (*bldN*) encoding a TCS regulator and a sigma factor, respectively, were overexpressed in the DEPA strain by 3.3- and 3.2-fold (Table 1).

A TCS response regulator gene (*SCLAV_p1081;* 3.5 fold) and the genes encoding anti-sigma regulatory factor

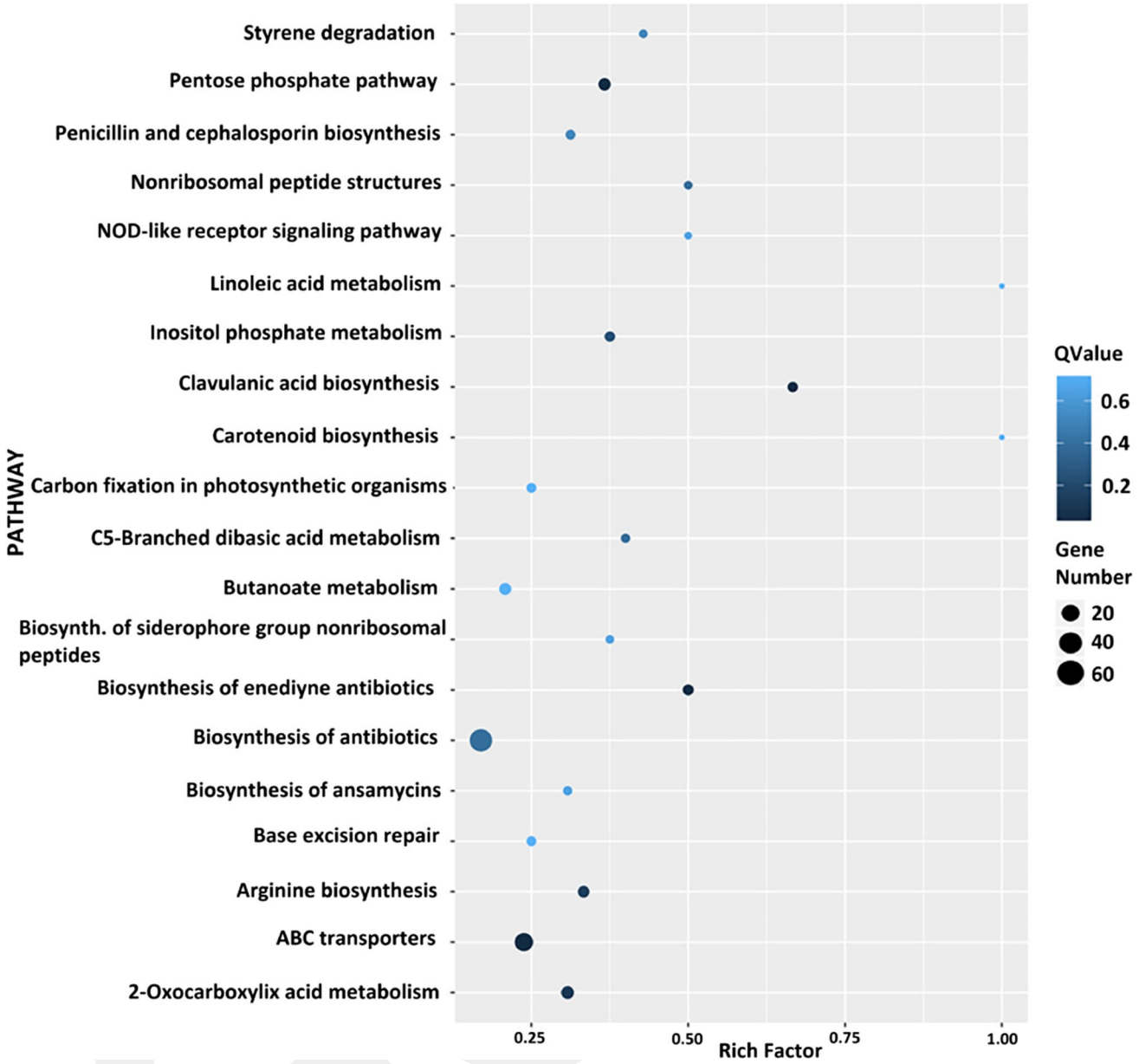


**Fig. 2** Bubble diagram of enrichment of KEGG pathway. The *X*-axis for the enrichment factor value and the *Y*-axis represents the pathway names. The size of the bubbles means the number of DEGs, while the color of the bubbles refers to Qvalue. The deeper the color represents the Qvalue is small, the smaller the value is, the more significant the enrichment results are

Ser/Thr protein kinase (*SCLAV_1592*) and a XRE family transcriptional regulator (*tpgCL1*, *SSCG_05675*, *SCLAV_1642)* were among the most upregulated genes with 2.6- to 2.7-fold increases. Lastly, the pyruvate, phosphate dikinase encoding gene *SCLAV_5441* (4.6-fold), and a DUF4255 domain-containing protein-encoding gene *SCLAV_1012* (4.5-fold) were also overexpressed in the industrial strain.

### The most significantly downregulated genes in the DEPA strain

The most significantly downregulated genes in *S. clavuligerus* DEPA are shown in Table 2 by excluding the genes encoding yet uncharacterized proteins. Mainly, genes located on plasmids showed decreased expression levels. For instance, *SCLAV_p0718* (putative monooxygenase) had the

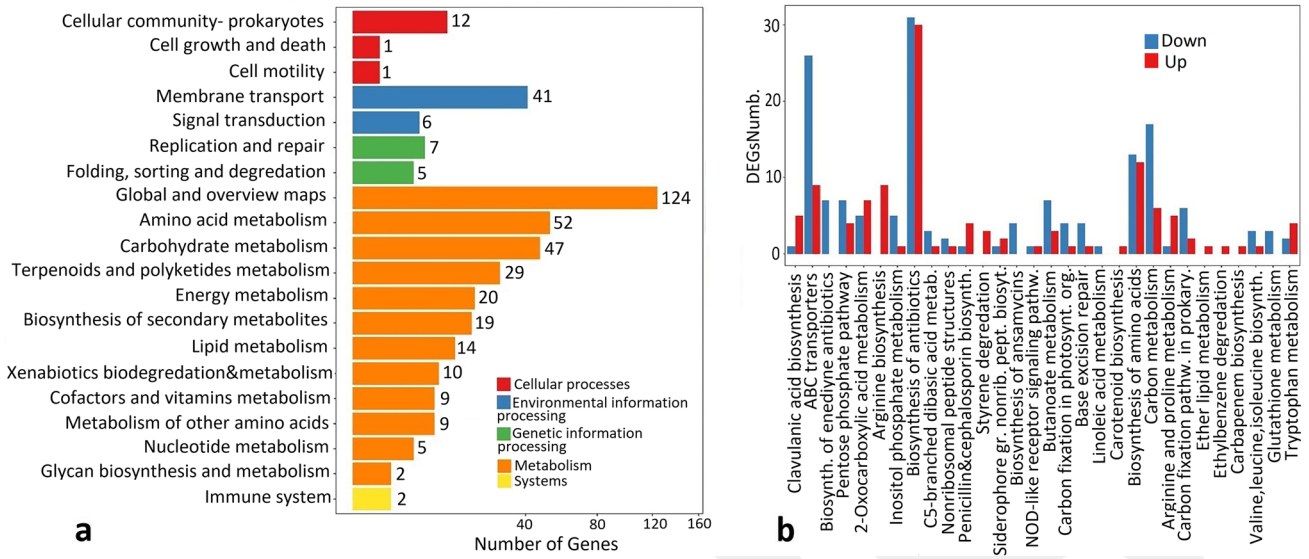


**Fig. 3** KEGG pathway classification analysis of significantly differentially expressed genes (**a**). X-axis represents number of genes significantly differentially expressed and Y-axis represents pathway classification represented by a different color according to the annotation. **b** Significantly up or downregulated gene (DEGs) number according to pathway classification. *Biosynth. of enediyne antibiotics*, Biosynthesis of enediyne antibiotics; *C5-branched dibasic acid metab.*, *C5-branched dibasic acid metabolism*; *Penicillin&cephalosporin biosynth.*, Penicillin&cephalosporin biosynthesis; *Siderophore gr. nonrib. pept. biosyt.*, Siderophore group non-ribosomal peptides biosynthesis; *NOD-like receptor signaling pathw.*, NOD-like receptor signaling pathway; *Carbon fixation in photosynt. org.*, Carbon fixation in photosynthetic organisms; *Carbon fixation pathw. in prokary.*, Carbon fixation pathways in prokaryotes; *Valine, leucine, isoleucine biosynth.*, Valine, leucine, isoleucine biosynthesis

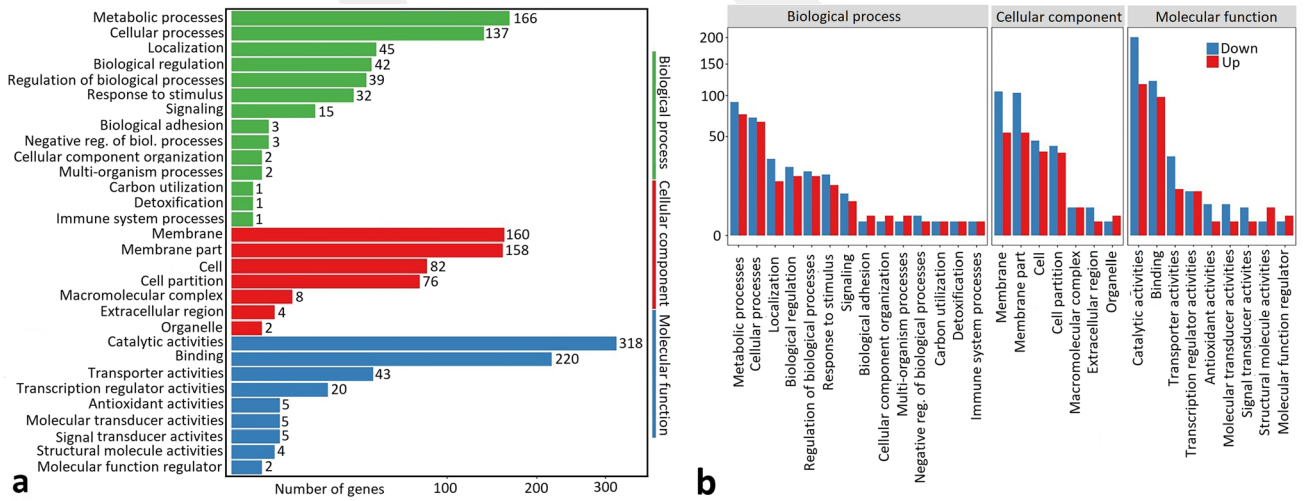


**Fig. 4** GO function classification map of significantly differentially expressed genes (**a**). Significantly up-/downregulated differential expression gene GO function classification map (**b**). *Negative reg. of biol. processes*, Negative regulation of biological processes

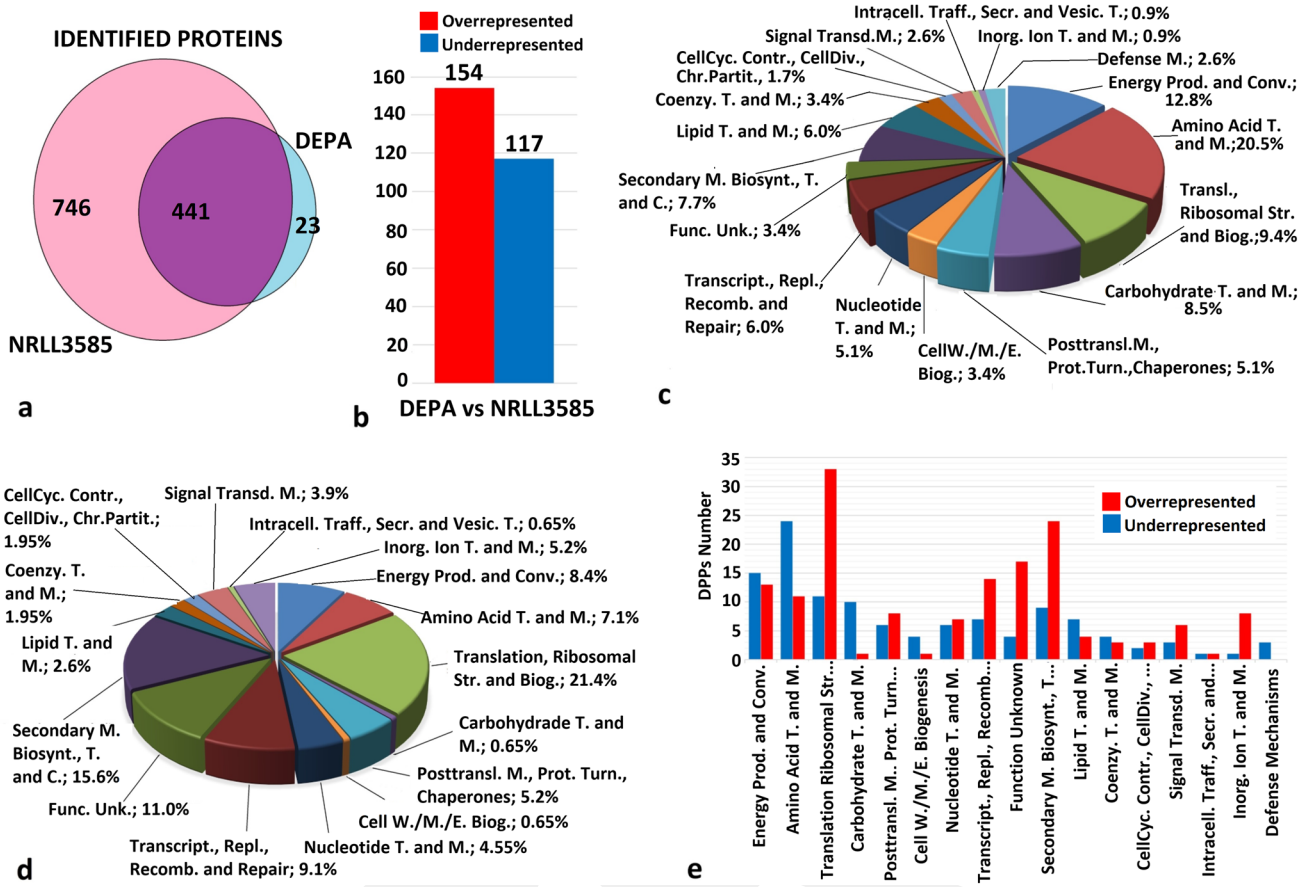


**Fig. 5** A Venn diagram showing common protein representation between *S. clavuligerus* NRRL3585 and strain DEPA (**a**). A bar graph showing statistics of DPPs in *S. clavuligerus* DEPA (**b**), and distribution of DPPs in DEPA strain identified by nanoLC-MS/MS analysis. **c** Overrepresented proteins, **d** underrepresented proteins, **e** comparison of the number of DPPs according to functional categories. *Intrecell. Traff. Secr. and Vesic. T.* Intracellular trafficking, secretion and vesicular transport Inorg; *Ion T. and M.* Inorganic ion transport and metabolism; *Defense M.* Defense Mechanisms; *Energy Prod. and Conv.* Energy production and conversion; *Amino Acid T. and M.* Amino acid transport and metabolism; *Transl., Ribosomal Str. and Biog.* Translation, ribosomal structure and biogenesis; *Carbohydrate T. and M.* Carbohydrate transport and metabolism; *Posttransl. M., Prot. Turn., Chaperones* Posttranlational modification, protein turnover, chaperones; *CellW./M./E. Biog* Cell wall/membrane/envelope biogenesis; *Nucleotide T. and M.* Nucleotide transport and metabolism; *Transcript., Repl., Recomb. and Repair* Transcription, replication, recombination and repair; *Func. Unk.* Function Unknown; *Secondary M. Biosynth., T and C.* Secondary metabolites biosynthesis, transport and catabolism; *Lipid T. and M.* Lipid transport and metabolism; *Coenzy. T. and M.* Coenzyme transport and metabolism; *CellCyc. Contr., CellDiv., Chr.Partit.* Cell cycle control, cell division, chromosome partitioning; *Signal Transd. M.* Signal transduction mechanisms

most noteworthy decrease in its expression within the *S. clavuligerus* DEPA transcriptome (13.3-fold). *SCLAV_p1278* (*moeE5*) was detected as 11.6-fold downregulated. *SCLAV_p0726* and *SCLAV_p0732* transcripts with 10.8- and 9.2-fold decrease encode cytochrome P450 proteins which take a role in producing secondary metabolites such as macrolides in *Streptomyces* sp. (Cho et al. 2019). *SCLAV_p0760* (transcriptional regulator) was detected at a 7.6-fold decreased level. *SCLAV_p0716* (ArsR-family; 7.1-fold), and *SCLAV_p0745* (MerR-family; 11.1-fold) are the other transcriptional factors that were also significantly downregulated. The last two genes are located in a terpene biosynthetic gene cluster (Hwang et al. 2019). *SCLAV_p0703-p0714* genes (except for *SCLAV_p0704* and *SCLAV_p0710*), which are located in the "other secondary metabolites biosynthetic gene cluster" (Hwang et al. 2019), were downregulated in a range of 7.4- to 9.6-fold.

*SCLAV_p0727* (for an alcohol dehydrogenase) was 12.6-fold downregulated. In addition, *SCLAV_p0750* (for a subtilisin-like protease) and *SCLAV_1039* (for a putative secreted subtilisin-like serine protease) genes were expressed at 12.7- and 7.7-fold decreased levels, respectively. *SCLAV_p0720* (*glpP*) coding for glycerol-3-phosphate responsive anti-terminator GlpP was 9.3-fold downregulated as well. It is also important to note that *sigE* and *sigF* genes (encoding RNA polymerase sigma factors SigE and SigF) are in this category with 2- and onefold downregulation, respectively (Table 2).

**Table 1** The most highly expressed genes in *S. clavuligerus* DEPA

| Uniprot Nr | Protein name | Gene | $log_2FC$ | *P* value |
|---|---|---|---|---|
| B5GV89 | CRISPR-associated endoribonuclease | *SCLAV_1861* | 7.661 | $7.72\times10^{-05}$ |
| B5GZD1 | Epimerase | *SCLAV_5662 (rmlC)* | 6.334 | $1.01\times10^{-05}$ |
| B5GQQ0 | Anti-sigma regulatory factor Ser/Thr protein kinase | *SCLAV_1592* | 6.088 | $5.26\times10^{-129}$ |
| B5GPH9 | Transcriptional regulator WhiB | *SCLAV_2475 (whiB)* | 5.746 | $1.16\times10^{-03}$ |
| D5SKD6 | TetR/AcrR family transcriptional regulator | *SCLAV_p0892* | 5.428 | 0 |
| E2Q0G7 | TetR/AcrR family transcriptional regulator | *SCLAV_5443* | 4.86 | 0 |
| E2Q8D9 | ABC-type multidrug transport system, ATPase and permease component | *SCLAV_0395* | 4.823 | $8.15\times10^{-239}$ |
| E2Q5V4 | Phage tail region protein | *SCLAV_0038* | 4.659 | $1.78\times10^{-61}$ |
| E2Q0G5 | Pyruvate phosphate dikinase (EC 2.7.9.1) | *SCLAV_5441 (pddK)* | 4.559 | 0 |
| B5H0W2 | DUF320 domain-containing protein (Chaplin protein) | *SCLAV_1012* | 4.485 | $6.33\times10^{-63}$ |
| E2PW36 | AmfT protein (amfT) class III lanthionine synthetase LanKC | *SCLAV_4953* | 4.270 | $2.14\times10^{-257}$ |
| E2Q5V5 | Putative phage tail sheath protein | *SCLAV_0039* | 4.075 | $1.42\times10^{-116}$ |
| E2PW37 | Membrane translocator (AmfB) | *SCLAV_4954* | 3.671 | $6.15\times10^{-108}$ |
| E2PY81 | Putative Mce family protein | *SCLAV_5185* | 3.575 | $7.82\times10^{-11}$ |
| D5SKD5 | TetR/AcrR family transcriptional regulator | *SCLAV_p0891* | 3.562 | $5.78\times10^{-134}$ |
| B5H0E0 | ABC-transporter integral membrane protein | *SCLAV_5183* | 3.475 | $1.10\times10^{-07}$ |
| D5SK74 | Transport permease protein (ABC type) | *SCLAV_p0831* | 3.452 | $7.12\times10^{-50}$ |
| B5GMR1 | Response regulator receiver protein | *SCLAV_p1081* | 3.448 | $2.70\times10^{-05}$ |
| B5GYV1 | Two component transcriptional regulator, LuxR family | *SCLAV_3699 (bldM)* | 3.310 | 0 |
| E2Q7L6 | AdsA/RNA polymerase sigma-70 factor | *SCLAV_2349 (adsA)* | 3.221 | $4.08\times10^{-79}$ |
| E2Q5W4 | Phage tail protein | *SCLAV_0048* | 3.166 | $1.29\times10^{-28}$ |
| B5GMZ5 | Daunorubicin resistance ABC transporter ATP-binding subunit | *SCLAV_p0830* | 2.985 | $1.35\times10^{-130}$ |
| B5H237 | XRE family transcriptional regulator | *TpgCL1* | 2.671 | 0 |
| B5H236 | XRE family transcriptional regulator | *SSCG_05675* | 2.615 | 0 |
| B5GQK1 | XRE family transcriptional regulator | *SCLAV_1642* | 2.606 | $8.86\times10^{-300}$ |

## Comparison of DEGs and DPPs in NRRL 3585 and DEPA strains based on functional categories

### DEGs and DPPs in relation to CA biosynthesis and CA precursors

The DEPA strain produces at least fourfold CA in SA cultures in relation to the 3585 strain (Figure S3 c,d). RNA analysis showed that *ceaS2*, *bls2*, *cas2*, *oat2*, *oppA1*, *cad*, and *orf14* genes were upregulated in the DEPA strain in relation to the 3585 strain. Concomitantly, their protein products were overrepresented in the DEPA strain as shown by nanoLC-MS/MS analyses (Table S2). The genes *claR*, (for a CSR gene of the CA cluster), *pah* (proclavaminate amidinohydrolase), and *cpe* (cephalosporin esterase, *orf12*) showed increased expression levels, but the encoded proteins ClaR, Pah, and Cpe were not overrepresented. *SCLAV_4184* (*orf16*) was the only gene downregulated in the CA cluster. However, it was not identified as an underrepresented protein either. Other genes in the CA cluster showed no expression differences in the DEPA strain (Table S2). Our RNA-seq analysis revealed in the DEPA strain upregulation of all arginine biosynthetic genes and also *argR*, encoding an arginine repressor gene. Among them, only ArgH was underrepresented at the proteome level. In relation to the G3P biosynt

hesis and glycerol metabolism, *pfk*A (*SCLAV_4332)* was threefold upregulated while the transcript levels of *pgk*, *pgi*, a different *pfk*A (*SCLAV_0508*), *fucA*, *fba, gap1, gap2, glpF2, glpK*, and *glpD1* were not significantly altered (Table S2).

### Clavam and alanylclavam biosynthesis

The genes located in the clavam biosynthetic cluster, namely, *cvm1-6* were significantly downregulated in the DEPA strain, while the transcript levels of *cvm7-cvm13, cvmH, cvmP,* and *cvmG* genes were not significantly altered. Nonetheless, there were no nanoLC-MS/MS data in regard to the remaining clavam biosynthesis-related enzymes. Regarding the expression of the CA paralog cluster genes, no significant gene expression differences were observed for *SCLAV_p1071, SCLAV_p1072*, *oat1, ceas1, pah1*, and *cvm7p* while *cvm6p*, *SCLAV_p1067*, *SCLAV_p1069*, and *SCLAV_p1070* genes were downregulated.

**Table 2** The most significantly downregulated genes in *S. clavuligerus* DEPA

| Uniprot number | Protein name | Gene | $log_2FC$ | *P* value |
|---|---|---|---|---|
| D5SJW2 | Putative monooxygenase | *SCLAV_p0718* | − 13.298 | 0 |
| D5SJZ4 | Subtilisin-like protease | *SCLAV_p0750* | − 12.707 | 0 |
| B5GNA3 | Alcohol dehydrogenase | *SCLAV_p0727 (adhA)* | − 12.551 | 0 |
| B5GNA5 | Putative endoribonuclease (YjgF) | *SCLAV_p0725* | − 11.687 | $5.52\times10^{-127}$ |
| D5SLH1 | NAD-dependent epimerase/dehydratase | *SCLAV_p1278 (moeE5)* | − 11.575 | 0 |
| B5GN83 | MerR-family transcriptional regulator | *SCLAV_p0745* | − 11.072 | $5.95\times10^{-184}$ |
| E2Q793 | Putative MutT-family protein | *SCLAV_0264* | − 10.824 | $2.92\times10^{-87}$ |
| D5SJX0 | Cytochrome P450 hydroxylase | *SCLAV_p0726* | − 10.812 | $1.10\times10^{-239}$ |
| B5GN81 | Chitin-binding protein | *SCLAV_p0747* | − 10.466 | $5.97\times10^{-83}$ |
| B5GNC8 | Carbonic anhydrase (EC 4.2.1.1) | *SCLAV_p0703* | − 9.609 | $1.2\times10^{-46}$ |
| D5SJW4 | Glycerol-3-phosphate responsive anti-terminator | *SCLAV_p0720 (glpP)* | − 9.305 | $7.75\times10^{-36}$ |
| D5SJX6 | Cytochrome P450 | *SCLAV_p0732* | − 9.201 | $8.78\times10^{-178}$ |
| B5GNC3 | Holo-acp synthase | *SCLAV_p0707* | − 9.116 | $2.33\times10^{-13}$ |
| B5GNB7 | Transketolase A subunit (EC 2.2.1.1) | *SCLAV_p0713* | − 9.076 | $4.48\times10^{-45}$ |
| B5GNB8 | Permease | *SCLAV_p0712* | − 8.775 | $4.35\times10^{-61}$ |
| D5SJV0 | ACPS domain-containing protein | *SCLAV_p0706* | − 8.399 | $1.13\times10^{-21}$ |
| D5SJV2 | tRNA synthetase | *SCLAV_p0708* | − 8.356 | $2.27\times10^{-29}$ |
| D5SJV5 | Condensation domain protein | *SCLAV_p0711* | − 8.265 | $3.02\times10^{-67}$ |
| B5GNC1 | Phytanoyl-CoA dioxygenase | *SCLAV_p0709* | − 7.744 | $1.12\times10^{-25}$ |
| E2PXJ5 | Putative secreted subtilisin-like serine protease | *SCLAV_1039* | − 7.744 | 0 |
| D5SJU9 | Putative transport protein | *SCLAV_p0705* | − 7.672 | $6.69\times10^{-19}$ |
| B5GN68 | Transcriptional regulator | *SCLAV_p0760* | − 7.639 | $7.97\times10^{-20}$ |
| B5GNB6 | Putative transketolase (Transketolase B subunit) | *SCLAV_p0714* | − 7.355 | $8.03\times10^{-16}$ |
| D5SJW0 | ArsR family transcriptional regulator | *SCLAV_p0716* | − 7.055 | $2.52\times10^{-38}$ |
| E2Q6B9 | RNA polymerase sigma factor SigF | *SCLAV_2170 (sigF)* | − 2.024 | 0 |
| B5GV44 | SigE family RNA polymerase sigma factor | *SCLAV_1910 (sigE)* | − 1.018 | $2.12\times10^{-16}$ |

### CephC biosynthesis

The DEPA strain produces in the order of 20–30% CephC in SA along the culture in relation to the wild-type strain (Figure S3 e,f). The transcripts of *lat*, *pcbAB*, and *pcbC* genes of the early steps of the CephC biosynthetic pathway were upregulated in the DEPA strain. The expression level of the *cmcT* gene that encodes a putative transporter for CephC was increased. No significant expression differences were observed for the remaining genes in the CephC gene cluster. Only Lat, PcbC, and CefF enzymes were overrepresented at the proteome level (Table S2). There was no significant differential expression recorded for the *ccaR* gene, a supercluster activator located in CephC gene cluster. Although some overexpressed/overrepresented genes/proteins in CephC cluster were recorded, a remarkable decrease in CephC production was observed in DEPA (Figure S3).

### Other secondary metabolite clusters

Production of holomycin was identical in strains DEPA and 3585 (Figure S3 i,j). Our data indicated that *hlmK* and *hlmM* genes, a biosynthetic and a regulatory gene of the holomycin gene cluster, respectively, were overexpressed in the DEPA strain. The expression levels of all other genes in this cluster are similar.

Tunicamycin was produced by the wild-type strain but was not detectable in the DEPA strain (Figure S3 g,h). However, the genes in the tunicamycin gene cluster were not differentially expressed in the DEPA strain in relation to the wild-type counterpart (Table S2). Only TunD (SCLAV_4184; glycosyl transferase group 1) was overrepresented in the DEPA strain in relation to the wild-type (Table S2).

We will analyze now the expression and representation of some genes/proteins located in different secondary metabolism gene clusters (SMC) in the DEPA strain in relation to the 3585 strain.

Several genes for type I, II, and III polyketide synthase (PKS) biosynthesis clusters were differentially expressed in *S. clavuligerus* DEPA. In SMCp16, which contains the genes for enediyne biosynthesis, and is located in plasmid pSCL4 (Medema et al. 2010), the *SCLAV_p1225* gene was upregulated, and its protein product, valine decarboxylase, was overrepresented. In cluster SCM5 (putatively for the

macrolide daptomycin formation) the *SCLAV_0452* and *SCLAV_0454* genes showed decreased expression levels. In contrast, *SCLAV_0464*, *SCLAV_0475*, *SCLAV_0476*, *SCLAV_0480*, and *SCLAV_0485* were upregulated. Also in SMC5, the *SCLAV_0471* (*avaA2*), a γ-butyrolactone biosynthesis protein did not display differential expression in spite of being this protein a target of the CagRS TCS system that partially regulates CA biosynthesis (Fu et al. 2019). Cluster SMC15 is putatively responsible for the polyketide spore pigment formation. In this cluster, the genes *SCLAV_4911* to *4921*, including *SCLAV_4913* (WhiEI homolog) and *SCLAV_4914* (Whi EII homolog) were upregulated in the DEPA strain.

The cluster SMC20 is composed of 34 genes, *SCLAV_5463* to *5497*, and is involved in the biosynthesis of a type III PKS. In this cluster, *SCLAV_5470, 5472, 5486, 5489, 5490, 5491, 5495,* and *_5496* were downregulated while only *SCLAV_5497* was upregulated in the DEPA strain (Table S2).

In relation to the biosynthesis of PKS-Nonribosomal Peptide Synthase (NRPS) type secondary metabolites, RNA-seq data revealed the upregulation of the *SCLAV_p1007* gene, located at SMCp11 (Medema et al. 2010). Compatible with this finding, its protein product was overrepresented in the DEPA strain by 3.0-fold, as was previously described (Ünsaldı et al. 2017). Cluster SMC13 (*SCLAV_ 4460* to *_4486*) is responsible for the formation of nanchangmycin (AbuSara et al. 2019), a product with a PKS-NRPS structure being *SCLAV_ 4466* the gene encoding the PKS-NRPS (Medema et al. 2010). In this cluster, *SCLAV_4468* and *SCLAV_4472* genes for a short-chain dehydrogenase and an amidase, respectively, were upregulated in the DEPA strain.

## Other metabolite DEGs/DPPs in *S. clavuligerus* DEPA

In the energy production and conversion category, pyruvate dehydrogenase (SCLAV_1401) and oxidoreductase (SCLAV_0466) were underrepresented, while BkdC (dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex) was overrepresented in the DEPA strain. SucB and SucC play a role in the Krebs cycle which are overrepresented in the DEPA strain.

In the amino acid transport and metabolism category, one of the underrepresented proteins, DapD (tetrahydro picolinate succinylase, *SCLAV_1115*), is responsible for the synthesis of LL-2,6-diaminopimelate (DAP) from (S)-tetrahydrodipicolinate via succinylase route. DAP is then converted to L-lysine by meso-diaminopimelate decarboxylase (LysA), one of the precursors for CephC biosynthesis. In addition, cysteine desulfurase (*SCLAV_1118*) is annotated to SufF that transfers sulfur from L-cysteine to yield L-alanine. Cysteine sulfurases are probably involved in cellular iron homeostasis and selenoprotein biosynthesis (Mihara and Esaki 2002). Underrepresentation of both proteins correlated well with decreased CephC production (Figure S3).

In the carbohydrate transport and metabolism section, *SCLAV_1133* showed a lower expression at the transcriptome level, and proteomics identified its corresponding transketolase as underrepresented. Mannose-1-phosphate guanylyltransferase (SCLAV_0620), transaldolase (SCLAV_1134), and oxppcycle protein (OpcA, *SCLAV_1136*) were also the underrepresented proteins with decreased expression levels in nanoLC-MS/MS (Table 3).

In the lipid transport and metabolism category, nanoLC-MS/MS analysis revealed decreased abundance of enoyl-[acyl-carrier-protein] reductase [NADH] (*SCLAV_1027*; FabI) in the DEPA strain.

In the posttranslational modification, protein turnover, chaperones category, and *SCLAV_2665 (grpE), SCLAV_2666 (dnaK)*, and *SCLAV_3692* (*groES*) transcripts were more abundant in DEPA strain. Concomitantly, their protein products, GrpE (HSP-70 cofactor), chaperone protein DnaK (HSP7), and 10 kDa chaperonin (GroES), were overrepresented. A 60 kDa chaperonin (GroEL) also displayed increased expression at the proteome level (Table 3).

Among the underrepresented proteins in the nucleotide transport and metabolism category, *SCLAV_0712*, CarB, which is related to L-arginine biosynthesis pathway, can be given as an example (Table 3).

A variety of transcriptional regulators were differentially expressed in DEPA strain such as the AraC-family transcriptional regulator (*SCLAV_1957*, AdpA) and nitrogen regulatory protein P-II (GlnB, *SCLAV_3545*) which were underrepresented in *S. clavuligerus* DEPA. In contrast BldD, a master repressor acting on AdpA, BldG, an anti-sigma factor antagonist (SCLAV_2542), and an anti-sigma factor (SCLAV_2541) were all abundantly represented. A TetR-family transcriptional regulator SCLAV_3146 and an iron repressor SCLAV_3415 were also regulator proteins overrepresented in *S. clavuligerus* DEPA.

### Function unknown

Several uncharacterized proteins were differentially expressed in the DEPA strain. Five of the overrepresented proteins in this category comprised *SCLAV_0317*, *SCLAV_4298*, *SCLAV_4781*, *SCLAV_5138*, and *SCLAV_5522*). Concomitantly, the *SCLAV_0317* gene provided an increased expression. *SCLAV_5522* and

**Table 3** Other metabolite DEGs/DPPs in *S. clavuligerus* DEPA

| Uniprot Nr | Protein name | Gene | $log_2$FC (RNA-seq) | $log_2$FC (nanoLC-MS/MS) |
|---|---|---|---|---|
| E2Q7L6 | AdsA/RNA polymerase sigma-70 factor | *SCLAV_2349 (adsa)* | 3.221 | – |
| E2Q9C7 | Putative gamma-butyrolactone biosynthesis protein | *SCLAV_0471 (avaA2)* | 0.290 | – |
| E2Q9B9 | Putative regulatory protein AfsR-like protein | *SCLAV_0463 (avaA1)* | −0.662 | – |
| E2Q9B8 | Gamma-butyrolactone biosynthesis protein | *SCLAV_0462 (afsR)* | 0.803 | – |
| E2Q0T7 | Pyruvate dehydrogenase E1 component | *SCLAV_1401* | −0.247 | −2.867 |
| B5GPS1 | Phosphoenolpyruvate carboxylase | *SCLAV_2248 (ppc)* | −0.200 | −1.656 |
| B5GPD0 | Aldehyde dehydrogenase | *SCLAV_2425* | −0.077 | −1.531 |
| B5GZV5 | Dehydrogenase | *SCLAV_2973 (paaZ)* | −0.243 | 2.380 |
| E2PWQ1 | Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex | *SCLAV_2979 (bkdC)* | 0.310 | 1.997 |
| E2Q1A2 | Monooxygenase FAD-binding protein | *SCLAV_3535* | −1.437 | 3.143 |
| E2Q5I4 | Dehydrogenase | *SCLAV_4123* | −0.486 | −1.765 |
| E2Q5I3 | Putative malate dehydrogenase | *SCLAV_4122* | −0.032 | 0.656 |
| B5H1L3 | Assimilatory ülfite reductase | *SCLAV_4899* | −0.142 | −1.117 |
| B5GSL5 | 4-hydroxyphenylpyruvate dioxygenase | *SCLAV_2046 (hpd)* | 0.355 | 1.374 |
| E2Q5C0 | Glutamate dehydrogenase | *SCLAV_2088* | −0.867 | −0.512 |
| B5GTG5 | Ornithine cyclodeaminase | *SCLAV_2697 (arcB)* | −0.354 | 0.744 |
| E2PYZ0 | Threonine synthase | *SCLAV_3250* | −1.437 | −0.757 |
| E2Q004 | Phosphoserine aminotransferase | *SCLAV_3352 (serC)* | −0.352 | −0.596 |
| B5GYY7 | Mannose-1-phosphate guanylyltransferase | *SCLAV_0620* | −0.610 | −2.326 |
| B5GYK3 | Transketolase | *SCLAV_1133* | −1.362 | −2.590 |
| E2PYJ0 | Transaldolase | *SCLAV_1134* | −0.930 | −1.488 |
| B5GYK0 | Oxppcycle protein | *SCLAV_1136 (opcA)* | −0.915 | −1.347 |
| E2QA08 | Ribonuclease J | *SCLAV_4635 (rnj)* | −0.288 | −1.008 |
| B5GX03 | Carbamoyl-phosphate synthase large chain | *SCLAV_0712 (carB)* | −0.354 | −1.452 |
| B5GV74 | Phosphoribosylaminoimidazolesuccinocarboxamide synthase | *SCLAV_1876* | 0.902 | 2.603 |
| B5GX28 | Adenylate kinase | *SCLAV_3658* | −0.334 | 0.698 |
| E2Q369 | GMP synthase [glutamine-hydrolyzing] | *SCLAV_3712 (guaA)* | −0.422 | 0.911 |
| B5H0X6 | Enoy-[acyl-carrier-protein] reductase [NADH] | SCLAV_1027 (*fabI*) | −0.209 | −1.641 |
| B5GWW9 | Probable transcriptional regulatory protein | *SCLAV_0752* | −0.387 | 0.792 |
| B5H1M7 | AraC-family transcriptional regulator | *SCLAV_1957 (adpA)* | 0.330 | −2.002 |
| B5GPA7 | TetR-family transcriptional regulator | *SCLAV_3146* | 0.996 | 1.240 |
| B5GP19 | Transcriptional factor regulator | *SCLAV_3224* | −0.344 | 1.073 |
| B5GNF7 | DtxR family transcriptional regulator, putative Mn-dependent transcriptional regulator (Iron repressor) | *SCLAV_3415* | −0.520 | 0.765 |
| A7DWG0 | AreB protein | *SCLAV_4500* | −0.435 | 0.328 |
| E2Q8Z2 | Nitrogen regulatory protein P-II | *SCLAV_4535 (glnB)* | 0.265 | −1.757 |
| E2PUF8 | Protein GrpE (HSP-70 cofactor) | *SCLAV_2665 (grpE)* | 1.584 | 0.674 |
| B5GTB1 | Chaperone protein DnaK (HSP70) | *SCLAV_2666 (dnaK)* | 1.643 | 1.045 |
| B5GYU4 | 10 kDa chaperonin (GroES protein) | *SCLAV_3692 (groES)* | 1.001 | 1.647 |
| B5GYU5 | 60 kDa chaperonin (GroEL protein) | *SCLAV_3693 (groEL)* | 0.852 | 0.830 |
| B5GPP5 | Anti-sigma factor (ATP binding protein) | *SCLAV_2541* | −0.241 | 0.944 |
| B5GPP6 | Anti-sigma factor antagonist | *SCLAV_2542 (bldG)* | −0.061 | 1.295 |
| B5GWZ7 | DNA-binding protein | *SCLAV_0719 (bldD)* | −0.389 | 1.112 |

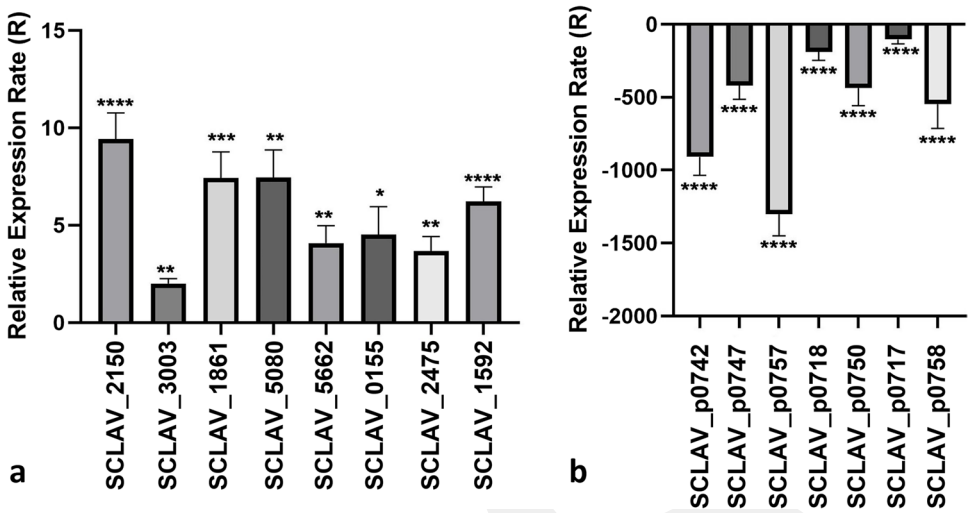


**Fig. 6** RT-qPCR validation of transcriptomic data. Relative expression ratios of overexpressed genes (**a**) and underexpressed genes (**b**) were obtained by RNA-seq experiments in *S. clavuligerus* DEPA compared to *S. clavuligerus* NRRL3585. Relative expression ratios for *S. clavuligerus* NRRL3585 strain was accepted as 1 (*: $P < 0.05$, **: $P < 0.01$, ****: $P < 0.0001$)

*SCLAV_2835* genes were shown by RNA-seq as downregulated, in contrast to their overrepresentation at the proteome level. Increased expression levels were recorded for DUF1775 domain-containing protein (*SCLAV_2835*) and DUF1416 domain-containing protein (*SCLAV_3194*). Putative nucleic acid–binding Zn-ribbon protein (*SCLAV_1535*), R3H domain-containing protein (*SCLAV_2904*), UPF0234 protein (*SCLAV_3612*), and Fe-S cluster assembly protein (SufD) were the remaining overexpressed proteins. R3H domain-containing protein is a predicted RNA binding protein (STRING). In addition, the STRING network annotates UPF0234 protein (*SCLAV_3612*) to ribosome-recycling factor (Frr) that plays a role in the release of ribosomes from mRNA at termination based on experimental data. SufD is a part of the SufBCD complex involved in Fe-S clusters' assembly under oxidative stress. It might contribute to iron uptake from chelators under iron limitation in *E. coli* (Outten et al. 2003).

Among the underrepresented proteins, no annotation is available for a secreted protein (*SCLAV_0126*). According to STRING predictions, the secreted tripeptidyl aminopeptidase (*SCLAV_0518*) was associated with hydrolases or acyltransferases alpha/beta hydrolase superfamily. Uncharacterized N-acetyltransferase (*SCLAV_1797*) displayed reduced expression in the DEPA strain at the proteome level.

## RT-qPCR validation of transcriptome data

A total of fifteen genes were selected for RT-qPCR analysis and transcriptome data validation. These included eight upregulated (*SCLAV_2150*, *SCLAV_3003, SCLAV_1861, SCLAV_5080, SCLAV_5662, SCLAV_0155, SCLAV_2475* and *SCLAV_1592*) and seven downregulated genes (*SCLAV_p0742, SCLAV_p0747, SCLAV_p0757, SCLAV_p0718, SCLAV_p0750, SCLAV_p0717* and *SCLAV_p0758* (Table S3). The expression levels of the selected DEGs accorded well with those obtained by transcriptomics (Fig. 6).

## Expression of potential regulons for β-lactam biosynthesis in S. clavuligerus DEPA

Hwang et al. (2019) predicted two prominent motifs by searching enriched -35 motifs in the promoter regions of the genes of β-lactam gene clusters and of the related ones as a potential regulon for β-lactam biosynthesis in *S. clavuligerus*. One was typical for *ceaS2* and *claR* in the CA gene cluster, *lat* and *blp* in the CephC gene cluster, *aceE* (pyruvate dehydrogenase gene), and *ilvC* (valine biosynthesis gene). The other motif was common in the promoter regions of *argG* and *argC* (arginine biosynthesis genes), *cvm7p* (a SARP-type CSR gene in paralog clavams gene cluster), *ceaS1*, and *res2* genes (clavam biosynthetic genes). These genes mostly showed a similar transcriptional regulation through the different growth stages of the strain. Therefore, upregulation of *ceaS2*, *claR*, *lat*, *argG*, and *argC* as DEGs and the presence of ArgG, ArgC, and AceE as DPPs fit well with the role of a specific regulon in the control of β-lactam biosynthesis in the DEPA strain. However, the remaining genes/proteins did not alter expression in *S. clavuligerus* DEPA. In the same study, 22 regulons regulated by different sigma factors were identified (Hwang et al. 2019). We screened all our DEGs/DPPs data for those potential regulons and found *oppA1*/OppA1 (regulon 8 and 13), *SCLAV_4781* (regulon 8), *SCLAV_4122* (regulon 3) *argR* (regulon 13), *SCLAV_p0713* (regulon 11 and 16), *SCLAV_2542* (regulon 17), *SCLAV_0131* (regulon 12), *SCLAV_3974* and *SCLAV_1752* (regulon 9).

### Correlation analysis between proteome and transcriptome data

A correlation coefficient value of $R = 0.27$ ($p = 1.99 \times 10^{-6}$) was obtained between the expression level of correlated 311 DEGs and DPPs, indicating a linear relationship within the range defined for bacteria (Chan 2003; de Sousa et al. 2009) (Figure S4).

## Discussion

We found an acceptable degree of consistency between the transcriptome and proteome data by correlation analysis. Integrative transcriptomic and proteomic analyses improve our understanding of the dynamic and complex interactions in biological systems (Kumar et al. 2016). Although many genes of the secondary metabolism category were determined at the transcriptome level, most of their corresponding proteins were not identified as DPPs by nanoLC-MS/MS analysis.

Our study does not show a clear correlation between the yield of CephC, tunicamycin or holomycin, and DEGs/DPPs present in the DEPA strain. However, this is not unusual and previous studies in *S. clavuligerus* reported a good yield of tunicamycins that was not related to increased expression of the tunicamycin genes and was explained by differences in the amounts of precursors (Martínez-Burgo et al. 2019). In the case of the clavams, Medema et al. (2011) described overexpression of genes in both the clavam and the paralogous gene cluster in an industrial strain. However, in the DEPA strain, we observed downregulation of many genes for clavams i.e. *SCLAV_p1067* to *_p1070* and *cvm6p* for the alanylclavam biosynthesis, and *cas1, cvm1* to *cvm6* for the clavam-specific late step. In addition, no expression differences were recorded for the paralog gene cluster in *S. clavuligerus* DEPA. The absence of clavams is probably an advantage for a CA overproducing strain since this allows the precursors to be specifically funnelled to CA formation (Paradkar 2013; Liras and Martín 2021).

The higher production of CA by *S. clavuligerus* DEPA was well correlated with the upregulation/overrepresentation of many genes and proteins. This is the case of the overrepresentation of the CeaS2, Cas2, Cad, Bls2, Oat2, OppA1, and Orf14 proteins that accorded well with our transcriptome data and with the higher CA production of the DEPA strain. Also, the higher CA production was correlated with better use of the CA precursors. Arginine is a precursor of CA and supplementation with arginine results in higher CA yields (Paradkar 2013). Arginine in *Streptomyces* is formed from glutamate in a cyclic pathway mediated by ArgJ (Rodríguez-García et al. 2000). All genes in the arginine pathway were overexpressed in the DEPA strain, as well as *oat2*, a gene located in the CA cluster. The *oat2* gene, encodes an ornithine acetyltransferase, that when disrupted decreases CA formation (de la Fuente et al. 2004). Although its exact role in CA formation has not been elucidated, Oat2 is a homolog to ArgJ (de la Fuente et al. 2004) and may increase the flow to arginine formation.

The primary carbon source in the SA medium is starch which is finally converted to glyceraldehyde 3-phosphate (G3P). In the DEPA strain, the phosphoglycerate kinase (Pgk) was underrepresented suggesting that the flow from G3P to pyruvate was lower, leading to G3P accumulation. The higher pool of G3P in the DEPA strain improves CA formation as shown previously by the disruption of the *gap1* gene for G3P degradation (Li and Townsend 2006).

Surprisingly none of the activators of the CA pathway, neither the main activator CcaR (Pérez-Llarena et al. 1997) nor the Orf22/Orf23 proteins, forming a CagRS two-component system, differs in our transcriptome or proteome data in relation to the wild-type strain. This is in contrast to what was previously reported for other *S. clavuligerus* CA overproducing industrial strains (Medema et al. 2011; Jnawali et al. 2008; Fu et al. 2019). Disruption of the CagRS system only partially blocks CA biosynthesis but exerts positive and negative effects on G3P and arginine biosynthesis, respectively, (Fu et al. 2019) indicating a complex effect on CA formation.

CcaR is a positive regulator of CephC and CA biosynthesis. RT-qPCR and microarray studies show that mutants deleted in CcaR do not produce CA at all, and the production of clavams, CephC, and holomycin decreases (Santamarta et al. 2011; Kurt et al. 2013; Álvarez-Álvarez et al. 2014). CcaR strongly activates the expression of the CA biosynthesis early genes (*ceaS2*, *bls2*, *cas2*, *pah2, oat2*, *oppA1*) and activates the expression of *claR,* a gene located in the CA cluster. Although *ccaR* expression was similar in both strains used in this study, upregulation of *claR* encoding a LysR type regulator was quite clear in the DEPA strain. ClaR exerts a positive effect on the expression of the late genes of the CA pathway (*cyp, oppA1*, and *cad*). However, in spite of being cluster-specific regulators, the lack of either CcaR or ClaR resulted in a global regulatory role on other regulons of antibiotic biosynthesis different from CA or CephC (Martínez-Burgo et al. 2015; Álvarez-Álvarez et al. 2014). For instance, *amfT* (for a class III lanthionine synthetase) is downregulated in the absence of ClaR (Martínez-Burgo et al. 2015); in our case, *amfT* is upregulated simultaneously with *claR* overexpression in the DEPA strain.

In *Streptomyces*, morphological development and antibiotic production are generally coincident and tightly controlled by a complex regulatory cascade in response to several environmental and physiological fluctuations. CSRs mediate the basic control of antibiotic biosynthesis but they are also under the control of higher-level regulators (Liu

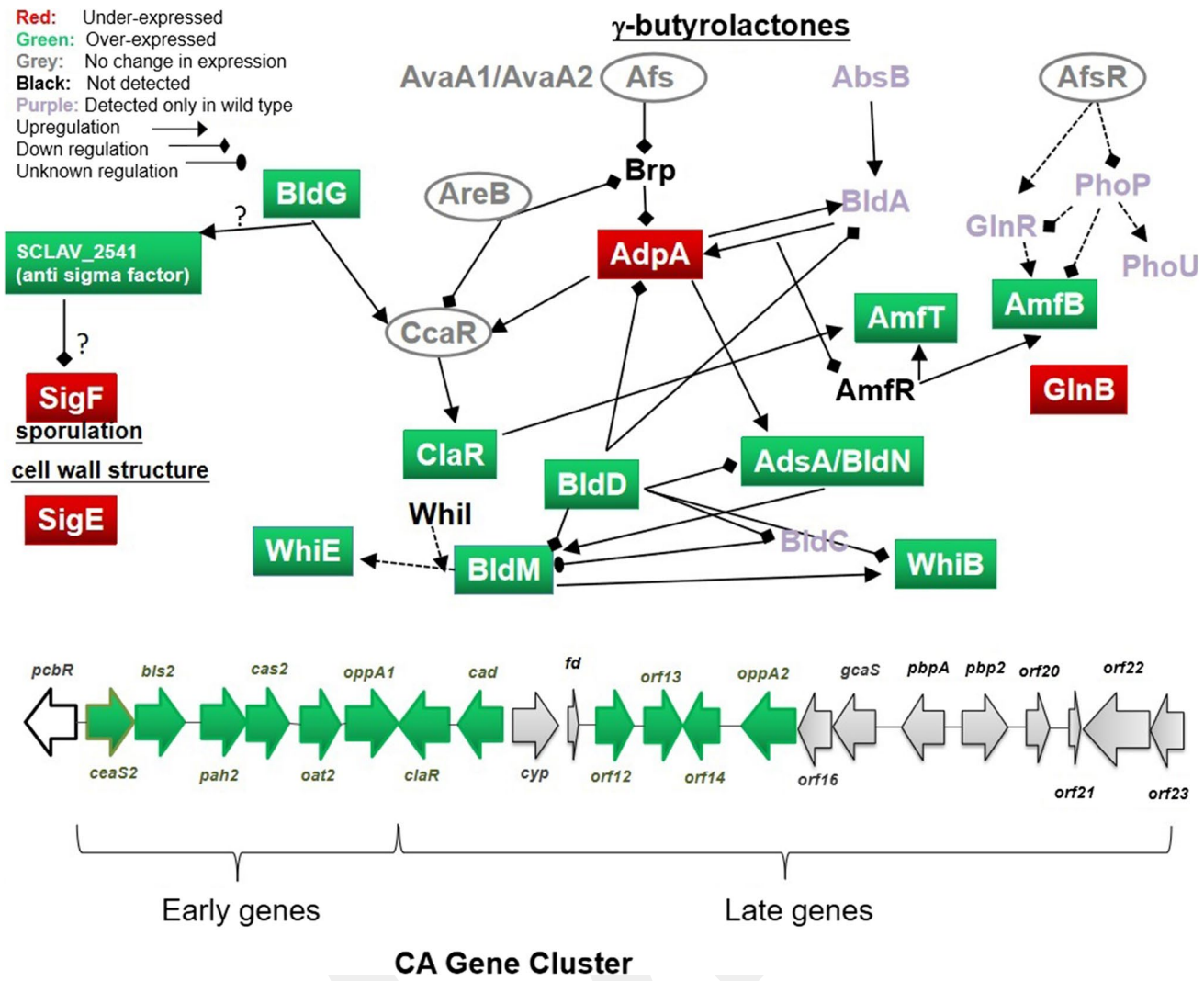


**Fig. 7** Regulatory cascade mechanism controlled by BldD and its relation to CA overproduction mediated by CcaR in *S. clavuligerus* DEPA. The results are obtained from previous reports (Martín et al. 2011; Bush et al. 2015; Ferguson et al. 2016; Ünsaldı et al. 2017) and from our study

et al. 2013). At the top of the regulatory cascade, BldD is a master developmental regulator which negatively controls the expression of other developmental regulatory genes such as *bldA*, *bldM*, *bldN/adsA, whiG*, and *whiB* (Bush et al. 2015). BldD is mainly known for its repressor function in developmental stages but may act as an activator in some *Streptomyces* strains (Yan et al. 2020). The regulatory network involving BldD is shown in Fig. 7. The lack of significant differences in expression of *ccaR* in *S. clavuligerus* DEPA may be due to a compensatory balance between the overexpression of BldG that activates CcaR (Bignell et al. 2005) and the underexpression of the *adpA* gene. This gene encodes a pleiotropic regulator, AdpA, controlling a regulon of genes for morphological differentiation and antibiotic production, including CA López-García et al. 2010; Higo et al. 2012); the low abundance of AdpA in the DEPA strain occurred most probably due to the BldD overrepresentation in this strain.

BldG is an anti-anti-sigma factor that controls *ccaR* expression and indirectly CA and CephC formation (Bignell et al. 2005). BldG and the protein encoded by the adjacent gene, an anti-sigma factor (homologous to *Bacillus* SpoIIA) were overrepresented in the DEPA strain. In *Bacillus*, SpoIIA represses SigF. We also showed downregulation of the *sigF* gene in *S. clavuligerus* DEPA suggesting that a cascade similar to that of *Bacillus* exists in *S clavuligerus.*

The higher level of BldM in the DEPA strain explains also the upregulation of *whiB* and *whiE.* BldM positively regulates the expression of *whiB*, a regulator that affects the formation of spores (Bush et al. 2017; Nah et al. 2021). In addition, BldM together with WhiI positively acts on *whiE* expression (Bush et al. 2015).

Overall, this is the first report showing BldD overrepresentation by an -omic study in an industrial CA overproducer *S. clavuligerus* strain. Our study also confirmed the interrelationship between several regulatory proteins such as BldD, BldG, BldM, BldN, and AdpA described by other authors (Bignell et al. 2005; Medema et al. 2011; Álvarez-Álvarez et al. 2014; Martínez-Burgo et al. 2015; Bush et al. 2015; Ferguson et al. 2016; Ünsaldı et al. 2017; Martínez-Burgo et al. 2019; Hwang et al. 2019; Yan et al. 2020; Nah et al. 2021). In addition, the late-step genes in the clavams biosynthetic cluster were downregulated considerably. It is also remarkable that arginine biosynthetic genes were highly upregulated while no difference was recorded for G3P biosynthesis-related genes/proteins. Our study confirmed that overexpression of CA requires the overrepresentation of many enzymes of the pathway and a good flow of the precursor arginine, but not the overexpression of G3P biosynthesis-related genes/proteins. The latter is being reported for the first time. The most highly expressed as well as significantly downregulated genes are also tabulated for the first time in our report. To further confirm these data and find out the potential new targets for rational strain improvement, we are initiating a parallel study using multiple time points along *S. clavuligerus* NRRL3585 and *S. clavuligerus* DEPA fermentation.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00253-022-12098-4.

**Acknowledgements** We acknowledge 2211-C Doctorate Scholarship Programme of The Scientific and Technological Research Council of Turkey (TUBITAK) for granting EÜ. We thank Caner Aktaş for his help in validation experiments. We also thank Çiğdem Otur for her contribution to bioassay experiments. We acknowledge Prof. Paloma Liras for her critical reading and valuable comments on our manuscript.

**Author contribution** GÖ designed and supervised the work and finalized the manuscript. AKK carried out the interpretation of data and prepared the manuscript draft. EÜ and SÖ performed the proteome experiments, and ŞAG performed the related bioinformatic analyses. GÇ was involved in RNA-seq analysis, RT-qPCR experiments, and statistical analyses. All authors read and approved the final manuscript.

**Funding** This work was supported by the Scientific and Technical Research Council of Turkey [TÜBITAK-KBAG, Grant No: 114Z648] and the Middle East Technical University Research Fund (BAP-01–08-2017–005).

**Data availability** All the data generated or analyzed during this study are included in this article (and its supplementary information files). The RNA-seq data were deposited in the SRA database under accession number SRP276705 while the proteome data are available at the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD032272.

## Declarations

**Ethics approval** This article does not contain any studies with human participants performed by any of the authors.

**Conflict of interest** The authors declare no competing interests.

## References

AbuSara NF, Piercey BM, Moore MA, Shaik AA, Nothias LF, Srivastava SK, Cruz-Morales P, Dorrestein PC, Barona-Gómez F, Tahlan K (2019) Comparative genomics and metabolomics analyses of clavulanic acid-producing *Streptomyces* species provides insight into specialized metabolism. Front Microbiol 10:2550. https://doi.org/10.3389/fmicb.2019.02550

Aharonowitz Y, Demain AL (1979) Nitrogen nutrition and regulation of cephalosporin production in *Streptomyces clavuligerus*. Can J Microbiol 25(1):61–67. https://doi.org/10.1139/m79-010

Álvarez-Álvarez R, Rodríguez-García A, Santamarta I, Pérez-Redondo R, Prieto-Domínguez A, Martínez-Burgo Y, Liras P (2014) Transcriptomic analysis of *Streptomyces clavuligerus* Δ*ccaR::tsr:* effects of the cephamycin C-clavulanic acid cluster regulator CcaR on global regulation. Microb Biotechnol 7(3):221–231. https://doi.org/10.1111/1751-7915.12109

Álvarez-Álvarez R, Martínez-Burgo Y, Rodríguez-García A, Liras P (2017) Discovering the potential of *S. clavuligerus* for bioactive compound production: cross-talk between the chromosome and the pSCL4 megaplasmid. BMC Genomics 18:1–13. https://doi.org/10.1186/s12864-017-4289-y

Baş L, Otur Ç, Kurt-Kızıldoğan A (2020) Enhanced tunicamycin biosynthesis in BldG overexpressed *Streptomyces clavuligerus*. Appl Biochem Microbiol 56:412–419. https://doi.org/10.1134/S0003683820040 02X

Bignell DR, Tahlan K, Colvin KR, Jensen SE, Leskiw BK (2005) Expression of *ccaR*, encoding the positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus*, is dependent on *bldG*. Antimicrob Agents Chemother 49(4):1529–1541. https://doi.org/10.1128/AAC.49.4.1529-1541.2005

Burton K (1968) Determination of DNA concentration with diphenylamine. Methods Enzymol 12:163–166. https://doi.org/10.1016/0076-6879(67)12127-7

Bush M, Tschowri N, Schlimpert S, Flärdh K, Buttner MJ (2015) c-di-GMP signalling and the regulation of developmental transitions in streptomycetes. Nat Rev Microbiol 13:749–760. https://doi.org/10.1038/nrmicro3546

Bush MJ, Chandra G, Findlay KC, Buttner MJ (2017) Multi-layered inhibition of *Streptomyces* development: BldO is a dedicated repressor of *whiB*. Mol Microbiol 104(5):700–711. https://doi.org/10.1111/mmi.13663

Byrum SD, Larson SK, Avaritt NL, Moreland LE, Mackintosh SG, Cheung WL, Tackett AJ (2013) Quantitative proteomics identifies activation of hallmark pathways of cancer in patient melanoma. J Proteomics Bioinform 6(3):043. https://doi.org/10.4172/jpb.1000260

Chan YH (2003) Biostatistics 104: correlational analysis. Singapore Med J 44(12):614–619

Cho MA, Han S, Lim YR, Kim V, Kim H, Kim D (2019) *Streptomyces* cytochrome P450 enzymes and their roles in the biosynthesis of macrolide therapeutic agents. Biomol Ther 27(2):127. https://doi.org/10.4062/biomolther.2018.183

De la Fuente A, Lorenzana LM, Martín JF, Liras P (2002) Mutants of *Streptomyces clavuligerus* with disruptions in different genes for clavulanic acid biosynthesis produce large amounts of holomycin: possible cross-regulation of two unrelated secondary metabolic pathways. J Bacteriol 184(23):6559–6565. https://doi.org/10.1128/JB.184.23.6559-6565.2002

De la Fuente A, Martin JF, Rodriguez-Garcia A, Liras P (2004) Two proteins with ornithine acetyltransferase activity show different

functions in *Streptomyces clavuligerus*: Oat2 modulates clavulanic acid biosynthesis in response to arginine. J Bacteriol 186(19):6501–6507. https://doi.org/10.1128/JB.186.19.6501-6507.2004

De Sousa AR, Penalva LO, Marcotte EM, Vogel C (2009) Global signatures of protein and mRNA expression levels. Mol Biosyst 5(12):1512–1526. https://doi.org/10.1039/B908315D

Deutsch EW, Bandeira N, Sharma V, Perez-Riverol Y, Carver JJ, Kundu DJ, García-Seisdedos D, Jarnuczak AF, Hewapathirana S, Pullman BS, Wertz J, Sun Z, Kawano S, Okuda S, Watanabe Y, Hermjakob H, MacLean B, MacCoss MJ, Zhu Y, Ishihama Y, Vizcaíno JA (2020) The ProteomeXchange consortium in 2020: enabling 'big data' approaches in proteomics. Nucleic Acids Res 48(D1):1145-D1152. https://doi.org/10.1093/nar/gkz984

Drawz SM, Bonomo RA (2010) Three decades of β-lactamase inhibitors. Clin Microbiol Rev 23(1):160–201. https://doi.org/10.1128/CMR.00037-09

Faurobert M, Pelpoir E, Chaïb J (2007) Phenol extraction of proteins for proteomic studies of recalcitrant plant tissues. Methods Mol Biol 355:9–14. https://doi.org/10.1385/1-59745-227-0:9

Ferguson NL, Peña-Castillo L, Moore MA, Bignell DR, Tahlan K (2016) Proteomics analysis of global regulatory cascades involved in clavulanic acid production and morphological development in *Streptomyces clavuligerus*. J Ind Microbiol Biotechnol 43(4):537–555. https://doi.org/10.1007/s10295-016-1733-y

Fu J, Qin R, Zong G, Liu C, Kang N, Zhong C, Cao G (2019) The CagRS two-component system regulates clavulanic acid metabolism via multiple pathways in *Streptomyces clavuligerus* F613–1. Front Microbiol 10:244. https://doi.org/10.3389/fmicb.2019.00244

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31(13):3784–3788. https://doi.org/10.1093/nar/gkg563

Higgens CE, Kastner RE (1971) Streptomyces clavuligerus sp., nov a β-lactam antibiotic producer. Int J Syst Evol Microbiol 21(4):326–331. https://doi.org/10.1099/00207713-21-4-326

Higo A, Hara H, Horinouchi S, Ohnishi Y (2012) Genome-wide distribution of AdpA, a global regulator for secondary metabolism and morphological differentiation in streptomyces, revealed the extent and complexity of the AdpA regulatory network. DNA Res 19:259–273. https://doi.org/10.1093/dnares/dss010

Hwang S, Lee N, Jeong Y, Lee Y, Kim W, Cho S, Palsson BO, Cho BK (2019) Primary transcriptome and translatome analysis determines transcriptional and translational regulatory elements encoded in the *Streptomyces clavuligerus* genome. Nucleic Acids Res 47(12):6114–6129. https://doi.org/10.1093/nar/gkz471

Jnawali HN, Oh TJ, Liou K, Park BC, Sohng JK (2008) A two-component regulatory system involved in clavulanic acid production. J Antibiot 61(11):651–659. https://doi.org/10.1038/ja.2008.92

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res 36:d480–d484. https://doi.org/10.1093/nar/gkm882

Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12(4):357–360. https://doi.org/10.1038/nmeth.3317

Kumar D, Bansal G, Narang A, Basak T, Abbas T, Dash D (2016) Integrating transcriptome and proteome profiling: strategies and applications. Proteomics 16(19):2533–2544. https://doi.org/10.1002/pmic.201600140

Kurt A, Álvarez-Álvarez R, Liras P, Özcengiz G (2013) Role of the *cmcH–ccaR* intergenic region and *ccaR* overexpression in cephamycin C biosynthesis in *Streptomyces clavuligerus*. Appl Microbiol Biotechnol 97(13):5869–5880. https://doi.org/10.1007/s00253-013-4721-4

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357–359. https://doi.org/10.1038/nmeth.1923

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12(1):323. https://doi.org/10.1186/1471-2105-12-323

Li R, Townsend CA (2006) Rational strain improvement for enhanced clavulanic acid production by genetic engineering of the glycolytic pathway in *Streptomyces clavuligerus*. Metab Eng 8(3):240–252. https://doi.org/10.1016/j.ymben.2006.01.003

Liras P, Martín JF (2021) *Streptomyces clavuligerus*: the omics era. J Ind Microbiol Biotechnol 48(9–10):1–18. https://doi.org/10.1093/jimb/kuab072

Liras P, Gomez-Escribano JP, Santamarta I (2008) Regulatory mechanisms controlling antibiotic production in *Streptomyces clavuligerus*. J Ind Microbiol Biotechnol 35(7):667–676. https://doi.org/10.1007/s10295-008-0351-8

Liu G, Chater KF, Chandra G, Niu G, Tan H (2013) Molecular regulation of antibiotic biosynthesis in Streptomyces. Microbiol Mol Biol Rev 77:112–143. https://doi.org/10.1128/MMBR.00054-12

López-García MT, Santamarta I, Liras P (2010) Morphological differentiation and clavulanic acid formation are affected in a *Streptomyces clavuligerus adpA*-deleted mutant. Microbiology 156(8):2354–2365. https://doi.org/10.1099/mic.0.035956-0

Martín JF, Sola-Landa A, Santos-Beneit F, Fernández-Martínez LT, Prieto C, Rodríguez-García A (2011) Cross-talk of global nutritional regulators in the control of primary and secondary metabolism in *Streptomyces*. J Microbial Biotechnol 4:165–174. https://doi.org/10.1111/j.1751-7915.2010.00235.x

Martínez-Burgo Y, Álvarez-Álvarez R, Rodríguez-García A, Liras P (2015) The pathway-specific regulator ClaR of *Streptomyces clavuligerus* has a global effect on the expression of genes for secondary metabolism and differentiation. Appl Environ Microbiol 81(19):6637–6648. https://doi.org/10.1128/AEM.00916-15

Martínez-Burgo Y, Santos-Aberturas J, Rodriguez-Garcia A, Barreales EG, Tormo JR, Truman AW, Reyes F, Aparicio JF, Liras P (2019) Activation of secondary metabolite gene clusters in *Streptomyces clavuligerus* by the PimM regulator of *Streptomyces natalensis*. Front Microbiol 10:580. https://doi.org/10.3389/fmicb.2019.00580

Medema MH, Trefzer A, Kovalchuk A, van den Berg M, Müller U, Heijne W, Wu L, Alam MT, Ronning CM, Nierman WC, Bovenberg RA, Breitling R, Takano E (2010) The sequence of a 1.8-Mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. Genome Biol Evol 2:212–224. https://doi.org/10.1093/gbe/evq013

Medema MH, Alam MT, Heijne WH, van den Berg MA, Müller U, Trefzer A, Bovenberg RA, Breitling R, Takano E (2011) Genome-wide gene expression changes in an industrial clavulanic acid overproduction strain of *Streptomyces clavuligerus*. Microb Biotechnol 4(2):300–305. https://doi.org/10.1111/j.1751-7915.2010.00226.x

Mihara H, Esaki N (2002) Bacterial cysteine desulfurases: their function and mechanisms. Appl Microbiol Biotechnol 60(1):12–23. https://doi.org/10.1007/s00253-002-1107-4

Nah HJ, Park J, Choi S, Kim ES (2021) WblA a global regulator of antibiotic biosynthesis in *Streptomyces*. J Ind Microbiol Biotechnol 48(3–4) https://doi.org/10.1093/jimb/kuab007

Nett M, Ikeda H, Moore BS (2009) Genomic basis for natural product biosynthetic diversity in the *Actinomycetes*. Nat Prod Rep 26(11):1362–1384. https://doi.org/10.1039/B817069J

Outten FW, Wood MJ, Muñoz FM, Storz G (2003) The SufE protein and the SufBCD complex enhance SufS cysteine desulfurase activity as part of a sulfur transfer pathway for Fe-S cluster assembly in *Escherichia coli*. J Biol Chem 278(46):45713–45719. https://doi.org/10.1074/jbc.M308004200

Özcengiz G, Demain AL (2013) Recent advances in the biosynthesis of penicillins, cephalosporins, and clavams and its regulation. Biotechnol Adv 31(2):287–311. https://doi.org/10.1016/j.biotechadv.2012.12.001

Özcengiz G, Okay S, Ünsaldı E, Taşkın B, Liras P, Piret J (2010) Homologous expression of aspartokinase (*ask*) gene in *Streptomyces clavuligerus* and its hom-deleted mutant: effects on cephamycin C production. Bioeng Bugs 1(3):191–197. https://doi.org/10.4161/bbug.1.3.11244

Paradkar A (2013) Clavulanic acid production by *Streptomyces clavuligerus*: biogenesis, regulation and strain improvement. J Antibiot 66(7):411–420. https://doi.org/10.1038/ja.2013.26

Pérez-Llarena FJ, Liras P, Rodriguez-Garcia A, Martin JF (1997) A regulatory gene (*ccaR*) required for cephamycin and clavulanic acid production in *Streptomyces clavuligerus*: amplification results in overproduction of both beta-lactam compounds. J Bacteriol 179(6):2053–2059. https://doi.org/10.1128/jb.179.6.2053-2059.1997

Perez-Riverol Y, Bai J, Bandla C, Hewapathirana S, García-Seisdedos D, Kamatchinathan S, Kundu D, Prakash A, Frericks-Zipper A, Eisenacher M, Walzer M, Wang S, Brazma A, Vizcaíno JA (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res 50(D1):D543–D552. https://doi.org/10.1093/nar/gkab1038

Pinilla L, Toro LF, Laing E, Alzate JF, Ríos-Estepa R (2019) Comparative transcriptome analysis of *Streptomyces clavuligerus* in response to favorable and restrictive nutritional conditions. Antibiotics 8(3):96. https://doi.org/10.3390/antibiotics8030096

Ramagli LS, Rodriguez LV (1985) Quantitation of microgram amounts of protein in two-dimensional polyacrylamide gel electrophoresis sample buffer. Electrophoresis 6(11):559–563. https://doi.org/10.1002/elps.1150061109

Rodríguez-García A, de la Fuente A, Pérez-Redondo R, Martín JF, Liras P (2000) Characterization and expression of the arginine biosynthesis gene cluster of Streptomyces clavuligerus. J Mol Microbiol Biotechnol 2(4):543–550

Santamarta I, Pérez-Redondo R, Lorenzana LM, Martìn JF, Liras P (2005) Different proteins bind to the butyrolactone receptor protein ARE sequence located upstream of the regulatory *ccaR* gene of *Streptomyces clavuligerus*. Mol Microbiol 56(3):824–835. https://doi.org/10.1111/j.1365-2958.2005.04581.x

Santamarta I, Lopez-García MT, Pérez-Redondo R, Koekman B, Martín JF, Liras P (2007) Connecting primary and secondary metabolism AreB an IclR-like protein, binds the ARE *ccaR* sequence of *S clavuligerus* and modulates leucine biosynthesis and cephamycin C and clavulanic acid production. Mol Microbiol 66(2):511e4 https://doi.org/10.1111/j.1365-2958.2007.05937.x

Santamarta I, López-García MT, Kurt A, Nárdiz N, Álvarez-Álvarez R, Pérez-Redondo R, Martin JF, Liras P (2011) Characterization of DNA-binding sequences for CcaR in the cephamycin-clavulanic acid supercluster of *Streptomyces clavuligerus*. Mol Microbiol 81(4):968–981. https://doi.org/10.1111/j.1365-2958.2011.07743.x

Song JY, Jensen SE, Lee KJ (2010) Clavulanic acid biosynthesis and genetic manipulation for its overproduction. Appl Microbiol Biotechnol 88(3):659–669. https://doi.org/10.1007/s00253-010-2801-2

Thorvaldsdóttir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14(2):178–192. https://doi.org/10.1093/bib/bbs017

Ünsaldı E, Kurt-Kızıldoğan A, Voigt B, Becher D, Özcengiz G (2017) Proteome-wide alterations in an industrial clavulanic acid producing strain of *Streptomyces clavuligerus*. Synth Syst Biotechnol 2(1):39–48. https://doi.org/10.1016/j.synbio.2016.10.003

Ünsaldı E, Kurt-Kızıldoğan A, Özcan S, Becher D, Voigt B, Aktaş C, Özcengiz G (2021) Proteomic analysis of a *hom*-disrupted, cephamycin C overproducing *Streptomyces clavuligerus*. Prot Pept Lett 28(2):205–220. https://doi.org/10.2174/0929866527666200723163655

Yan H, Lu X, Sun D, Zhuang S, Chen Q, Chen Z, Li J, Wen Y (2020) BldD, a master developmental repressor, activates antibiotic production in two *Streptomyces* species. Mol Microbiol 113(1):123–142. https://doi.org/10.1111/mmi.14405