

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/138641> holds various files of this Leiden University dissertation.

Author: Du, C.

Title: Multi-omics studies of the control of growth and antibiotic production of streptomyces

Issue Date: 2020-12-09

Unravelling the response of *Streptomyces roseifaciens* to challenge with small molecules by genome-wide proteomics

Chao Du, Somayah Elsayed, Marta del Rio, Bogdan I. Florea, Jos Raaijmakers, Gilles P. van Wezel

Manuscript in preparation

Abstract

Bacteria of the order of the Actinomycetales are the most prominent antibiotic producers. Of these, members of the genus *Streptomyces* produce half of all antibiotics used in the clinic today. To unlock the full potential of these bacteria in terms of natural product biosynthesis, eliciting strategies need to be developed that mimic the environmental conditions, to activate silent biosynthetic gene clusters (BGCs). Here, we analysed the response of the gifted antibiotic producer *Streptomyces roseifaciens* to challenge with small molecules, with the aim to find elicitors of the expression of cryptic BGCs. The bioactivity changes were correlated to the proteome level shifts using quantitative proteomics. Hydroxycoumarin was identified as an efficient elicitor, which may be explained by changes in energy metabolism. Surprisingly, despite their major differences on antimicrobial activity, only minor effects were seen in protein profiles when *S. roseifaciens* was exposed to jasmonic acid, chitosan, benzoic acid, or N-acetylglucosamine. A time series proteomics study on the immediate response of *S. roseifaciens* to jasmonic acid (JA) identified a BGC designated *jar* and a gene encodes an MFS transporter as major determinants for the JA response. The *jar* cluster directly correlated to the antimicrobial activity of *S. roseifaciens*, as shown by mutational analysis.

Introduction

Streptomycetes are soil-dwelling microorganisms with a complex multicellular life cycle (Claessen *et al.*, 2014). These bacteria are a major source of clinically approved drugs, such as antibiotics and anticancer compounds (Hopwood, 2007, Newman & Cragg, 2016, Barka *et al.*, 2016). As antibiotics are being widely used across the globe, the increasing antibiotic-resistant bacteria are becoming a severe problem. The failure to identify novel bioactive molecules is thereby a huge bottleneck in the battle against resistant microbes. Nowadays, the brute force antibiotic discovery gradually comes to an end as the easy to find compounds have been mostly discovered and synthetic antibiotics fail to meet the needs (Kolter & van Wezel, 2016). Advances in genomics showed that the potential of Actinobacteria as producers of bioactive natural products may have been grossly underestimated (Hopwood, 2007). Scientists now have the task to unleash the full producing capacity of these bacteria.

In bacteria, the biosynthetic genes responsible for the production of specific NPs are typically clustered, thus forming so-called biosynthetic gene clusters (BGCs). Bioinformatics algorithms based on the current understanding of BGCs often gives about 20 to 50 potential BGCs of different kinds from a single *Streptomyces* genome, many of which are silent or cryptic under routine laboratory conditions. One way to activate such silent BGCs is genetic manipulation of the regulatory network that controls the BGC, for example by over-expression of pathway-specific activators (Sanchez *et al.*, 2010, van der Heul *et al.*, 2018). Another method that is often used when prior information is not available, is to change the external and/or internal biological factors by fluctuating the growth environment (Blin *et al.*, 2017, Rutledge & Challis, 2015, Zhu *et al.*, 2014a). These so-called eliciting strategies aim to mimic the environmental conditions in the laboratory, where the activating factors of antibiotic production is often missing (van der Meij *et al.*, 2017).

There is a strong correlation between the production level of a certain NP and the expression of its cognate BGC (Du & van Wezel, 2018). Therefore, once fluctuation of NP production has been achieved, samples from different conditions can be compared using metabolomics. Quantitative proteomics or transcriptomics allow correlation of a specific BGC to its cognate bioactive compound (Gubbens *et al.*, 2012, Wu *et al.*, 2016a). This allows efficient dereplication of known molecules and gene clusters. Proteomics, as a very important method that allows to access the cellular machinery that directly participates in most biological process, has also evolved to have better coverage and accuracy during the years, and is becoming a very attractive tool in NP research (Du & van Wezel, 2018, Schubert *et al.*, 2017).

Since a large part of antibiotics is produced by soil dwelling bacteria of the genus *Streptomyces*, an important pathway of activating the cryptic antibiotic resources is to investigate the interactions between plant and microbe. We know that the ability of producing compounds like jasmonic acid, salicylic acid, and auxin is important for the plants to modulate their surrounding microbiome. This modulation is very important for pathogen resistance of the plants (van der Meij *et al.*, 2017, Xiao *et al.*, 2002, Gopalakrishnan *et al.*, 2011, Carvalhais *et al.*, 2015, Lebeis *et al.*, 2015). Recently, researchers proposed the *cry for help hypothesis*, which entails that plants recruit beneficial Actinobacteria and activate their antibiotic production by releasing stress-related molecules as exudates upon challenge by pests (van der Meij *et al.*, 2017). Recent studies in our laboratory have shown that *Streptomyces roseifaciens* responds to the plant hormone jasmonic acid by enhancing antibiotic production. *S. roseifaciens* is a very gifted strain in terms of NP biosynthesis, which originates from the Qinling mountains in China (van der Aart *et al.*, 2019). The *S. roseifaciens* genome contains 55 putative BGCs for small molecules as predicted by antiSMASH (Blin *et al.*, 2013). Indeed, the strain produces a wide range of NPs, including antibiotics with efficacy against multiple Gram-positive and Gram-negative multi-drug resistant pathogens (Wu *et al.*, 2016c, Zhu *et al.*, 2014b) and other NPs such as isocoumarins, prodiginines, acetyltryptamine, ferverulin and pyranonaphthoquinones (Wu *et al.*, 2016c, Wu *et al.*, 2017).

In this chapter, the response of *S. roseifaciens* to a range of small molecules are shown. A set of 12 small molecules was selected based on their different properties, including plant hormones, oligosaccharides and pathogen-related compounds. Quantitative proteomics was performed to obtain insights into the changes in the protein profiles, and correlate these to metabolic information of the NPs produced. The response to the 12 compounds could be clustered in three main groups, of which hydroxycoumarin stood out as a single group. The response to jasmonic acid (JA) was worked out more, and a 4 h time-course experiment following induction by JA revealed a JA-repressed gene cluster (*jar* cluster) and a JA-activated MFS-type transporter.

Results and Discussion

Small molecules selected for proteomics study of the response of *S. roseifaciens*

In order to explore the potential of *S. roseifaciens* as a natural product producer, a set of 12 different small molecules was selected to analyse their potential as elicitors of cryptic BGCs (Table 1). One major group of small molecules tested were plant hormone compounds, including jasmonic acid (JA), salicylic acid (SA), indol-3-acetic acid (IAA) and its synthetic analogue naphthaleneacetic acid (NAA). Of these, JA elicits the production of antibiotics in *S. roseifaciens* (van der Meij et al. unpublished data). The plant hormones IAA and NAA are auxins that induce plant cell elongation and cell division (Hopkins, 1999). Some *Streptomyces* species are also able to produce auxins, which might have a plant growth-promoting effect (Sadeghi et al., 2012, Viaene et al., 2016). A second group of small molecules consisted of hydroxycoumarin (HyC), N-acetylglucosamine (GlcNAc), cellobiose (CB), chitosan (CS), cinnamic acid (CA), ferulic acid (FA), benzoic acid (BA), and butyrate (BR). These compounds are related with cell components of other organisms or metabolites of those components. Of these, GlcNAc was previously shown to activate silent BGCs in *Streptomyces*, via the inactivation of the global antibiotic repressor DasR (Craig et al., 2012, Rigali et al., 2008, van Wezel et al., 2009). CS, which consists of both GlcNAc and GlcN units, is an incomplete deacetylation product of chitin. It can induce defence-related responses in plants (Hadwiger, 2013). CB is a hydrolysate of the highly abundant natural polymer cellulose, and can be efficiently used by *Streptomyces* (Colson et al., 2008, Liao et al., 2014). CA is a precursor of a large number of plant substances. FA and BA are derivatives of CA. All three molecules are metabolized by *Streptomyces* species (Sutherland et al., 1981, Sutherland et al., 1983).

Antimicrobial activity can be both induced and suppressed by the selected small molecules

To assess the actual effects of these selected small molecules, antimicrobial tests were done. *S. roseifaciens* spores were grown as spots on MM with the different compounds present in the medium, followed by overlay with *E. coli* strain ASD19 (Figure 1). This revealed major differences, whereby GlcNAc, CB, BR and CS decreased the total antibiotic activity of *S. roseifaciens*, while most other compounds elicited the bioactivity, except SA which had no noticeable effect. Interestingly, the compounds with the most distinct effect, namely CS, GlcNAc, JA and BA, clustered together in the proteomic response. As discussed later in this Chapter, in fact we did not see a clear correlation between the full proteome-level response and the antimicrobial activities.

Table 1. The small molecules used as putative elicitors in this study.

Abbreviation	Compound name	Notes
JA	Jasmonic acid	Plant hormone of jasmonate class, involved in the stress response and known as transducers of plant secondary metabolite production (Farmer et al., 2003).
SA	Salicylic acid	Plant hormone involved in response to stress and development (Jin et al., 2017, Ahmad et al., 2019).
IAA	Indole-3-acetic acid	Plant hormone of auxin class, can induce plant cell elongation and cell division (Hopkins, 1999).
NAA	Naphthalene acetic acid	Synthetic plant hormone in the auxin family (Hopkins, 1999).
HyC	Hydroxycoumarin	Fungal metabolite relating to plant cell damage (Bocks, 1967, Brown, 1986, Liu et al., 2009), mimicking fungal presence.
GlcNAc	N-acetylglucosamine	Involved in nutrient signalling of <i>Streptomyces</i> , related to regulation of antibiotics production (Rigali et al., 2008).
CB	Cellobiose	Disaccharide from the hydrolysis of cellulose, induces expression of cellulases in many fungi (Ilmén et al., 1997).
CS	Chitosan	Chitin deacetylation product.
CA	Cinnamic acid	Precursor for the synthesis of a huge number of plant substances, has antimicrobial activity (Ferreira & Teixeira, 2003).
FA	Ferulic acid	Product of cinnamic acid metabolism, abundant in plant cell wall, can be released by wounding. Induces <i>vir</i> genes of <i>Agrobacterium tumefaciens</i> (Heath et al., 1995, Kalogeraki et al., 1999).
BA	Benzoic acid	Product of cinnamic acid metabolism, rich in many crops, food preservative (del Olmo et al., 2017).
BR	Butyrate	End product of bacterial anaerobic fermentation, also a building block of NPs produced by <i>Streptomyces</i> (Omura et al., 1983, Reynolds et al., 1988).

Proteome level responses to small molecules

To obtain biomass for proteomics experiments, *S. roseifaciens* was pre-cultured on MM agar plates covered with a cellophane disc for 48 h at 30°C. After this, the cellophane disk with mycelial biomass was moved to a fresh MM agar plate containing one of the small molecules at a concentration of 500 µM or same amount of solvent as control. Then the mycelia were grown for another 48 h and harvested for protein extraction and proteomic measurement. Triplicate experiments were performed for each of the 13 growth conditions (12 small molecules + control), giving a total of 39 samples.

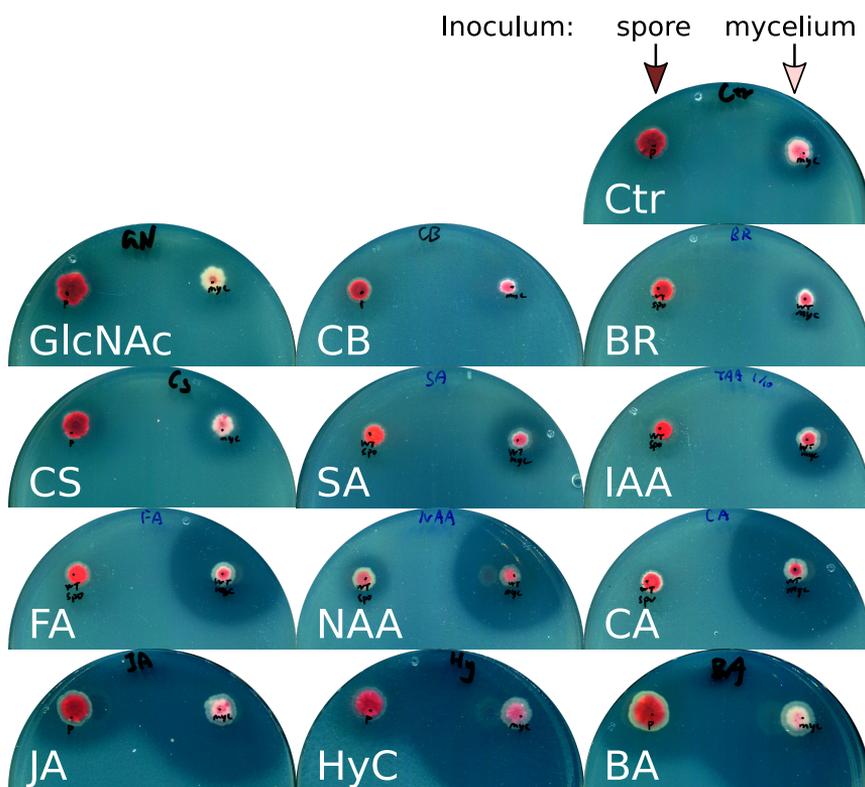


Figure 1. Effect of small molecules on the antimicrobial activity of *S. roseifaciens*.

Spots of spore (left) and mycelium (right) were grown on MM agar supplemented with 1% glycerol and 0.5% mannitol. Small molecules were added to 500 μ M in the media before inoculation. *E. coli* ASD19 was used as indicator strain, and overlay was done on the fifth day after inoculation. Pictures were taken 16 h after the overlay. Ctr for control.

Pearson's correlation coefficients were calculated to show the differential protein levels between different samples, while at the same time allowing general assessment of the quality of the data. A cut-off score of 0.95 was maintained for the correlation within triplicates. As a result, IAA replicate #1, FA replicate #2, CA replicate #1 and NAA replicate #1 were removed from the data for their low correlation with the other two replicates. The data normalization and quantification were done again using the rest of samples, ensuring best data quality. Subsequently, single quantifications among replicates were removed. In total, 1,897 proteins could be quantified in at least one condition. The number of quantified proteins in each sample ranged from 1,507 to 1,739, with an average of 1,638. Of the quantified proteins, 1,242 were shared by all 13 conditions (two or more identifications among replicates). The matrix of correlation coefficients revealed that hydroxycoumarin (HyC) samples had the most distinct proteome

profile as compared to samples obtained from growth on other small molecules (Figure 2 blue box), with a correlation coefficient < 0.9 for the comparison. The correlation matrix also showed that protein profiles were very similar for samples obtained from MM with either JA, GlcNAc, CS, or BA, with a correlation coefficient > 0.95 (Figure 2 red box). Principal component analysis (PCA) was done, and three clusters were found amongst all conditions (Figure 3). Samples treated with auxins (IAA, NAA), SA, CA and its derivative FA, CB, or BR clustered together with the control, which indicates that these compounds did not trigger a global change in gene expression in *S. roseifaciens*. As suggested by the high correlation, samples obtained from mycelia grown on JA, GlcNAc, chitosan or BA formed another closely related cluster, well separated from the control and related samples. Samples obtained from MM with hydroxycoumarin (HyC) were remote from these two large clusters. This indicates that HyC elicits a distinct systematic response as compared to the other compounds and the control.

4

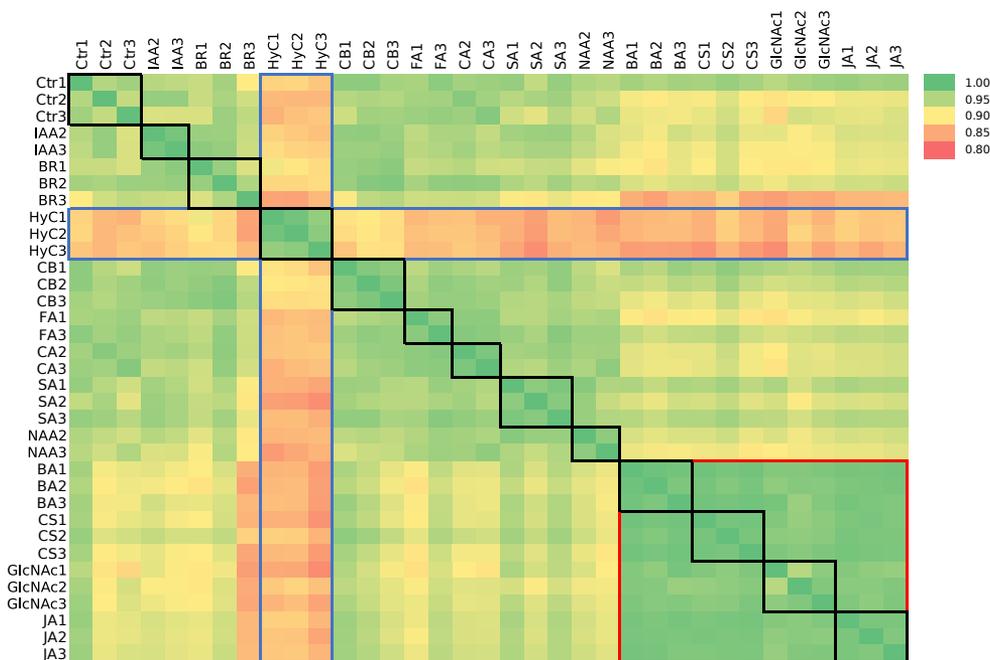


Figure 2. Correlations between protein profiles in response to small molecules.

Replicates are indicated using black borders (three replicates analysed, samples that failed to comply with statistical threshold not shown). Hydroxycoumarin-treated samples are indicated with a blue box. The four compounds giving highly similar responses are highlighted with a red box. Ctr for control.

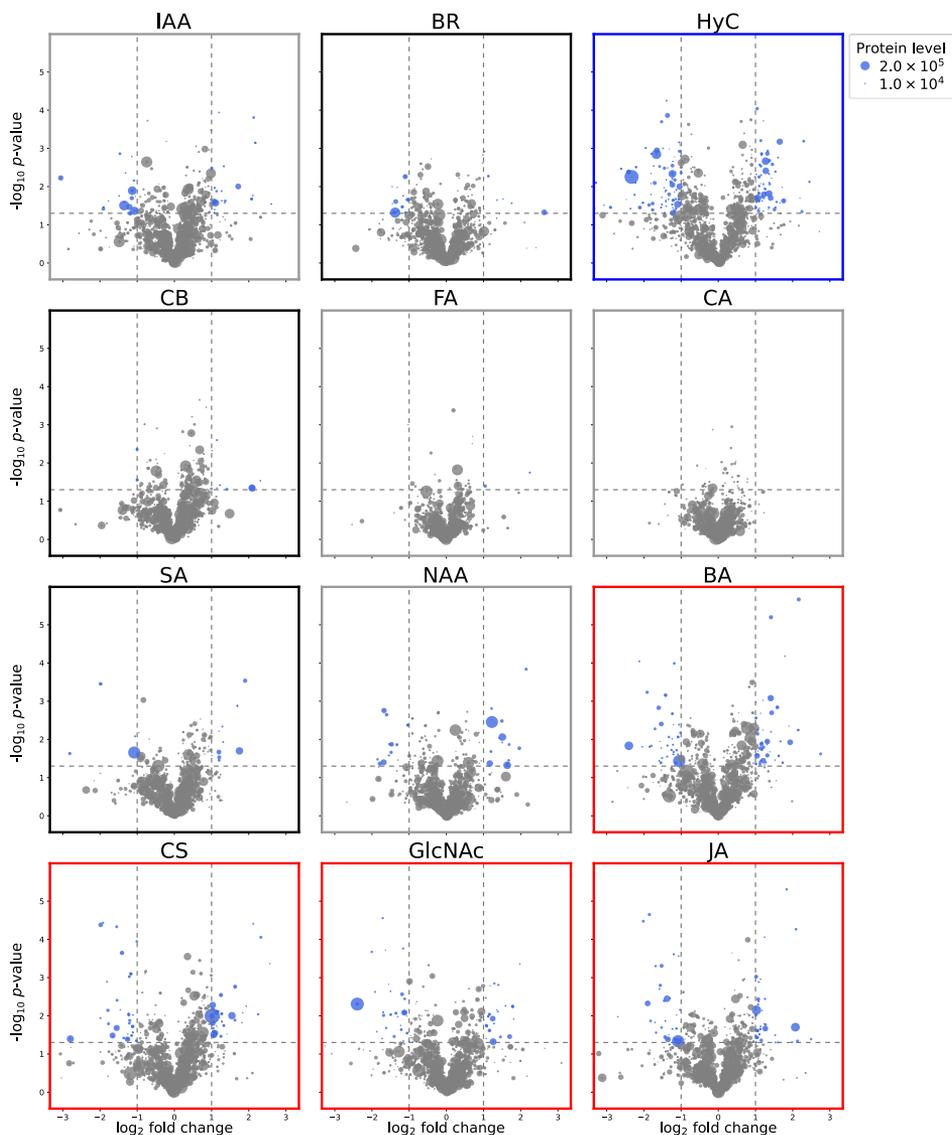


Figure 4. Volcano plots of DEPs following addition of different compounds. The significantly changed (fold change ≥ 2 , p -value < 0.05) proteins are indicated by blue dots, other proteins are shown as grey dots. The size of each dot represents the higher average protein level of control and test condition. HyC treatment is indicated by blue bordered plot. Red bordered plots indicate the four conditions that clustered closely (Figure 3). Grey bordered plots (IAA, NAA, FA and CA) are the treatments with one sample removed from the analysis, which led to higher p -values (lower $-\log_{10} p$ -value). The blue border plot corresponds to HyC samples, red border plots to the four clustered molecules (BA, CS, JA, and GlcNAc), as shown in Figure 3.

In order to get an overview of the function of DEPs, gene ontology (GO) enrichment analysis was performed for the DEPs of all treatments. For each treatment, three sets of proteins were tested for enrichment the GO aspect “biological process”, they were up-regulated, down-regulated and combined DEPs. All over-represented GO classes which were the lowest in the GO hierarchy were listed in Table 3. For the HyC treated samples, proteins related to “ATP synthesis coupled proton transport” were significantly overrepresented (corrected p -value < 0.05) in the extracted DEPs, whereby three of the five proteins belonging to this category were significantly up-regulated. This suggests that energy metabolism was affected by HyC, which may explain the fundamental changes in the overall proteomic shift shown in the correlation and PCA graph (Figure 2, Figure 3). Alpha-L-glutamate ligase SC03_2012 was among the most up-regulated DEPs following HyC treatment ($\log_2^{\text{fold change}} = 2.45$). This is a RimK-family ATP hydrolase that likely plays a role in the biosynthesis of small molecules (Zhao *et al.*, 2013). The two paralogues of pyruvate dehydrogenase E1 component (SC01_0222 and SC04_0475) were down-regulated when grown on HyC, with \log_2 fold changes of -3.29 and -0.76, respectively. The down-regulation of the key metabolic enzyme pyruvate dehydrogenase suggests changes in central metabolism.

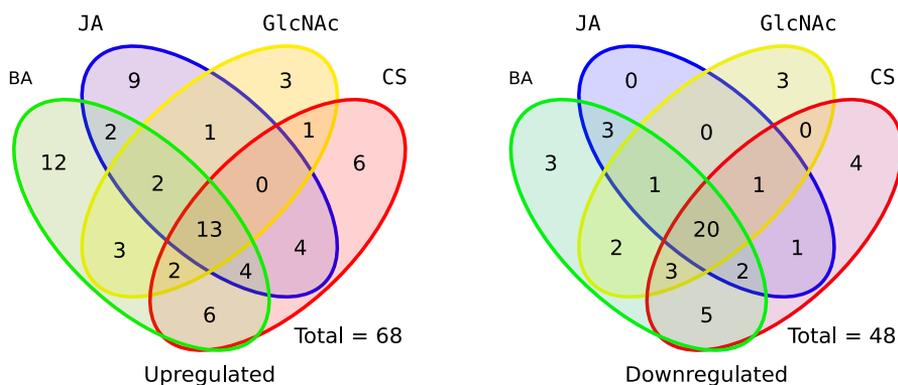


Figure 5. Venn diagram showing DEPs shared between four conditions that induced similar responses in *S. roseifaciens*. About 20% and 40% proteins were commonly up- and down-regulated, respectively, indicating a larger heterogeneity in the up-regulated proteins than down-regulated proteins.

Within the cluster of samples treated with BA, JA, CS, and GlcNAc, 13 proteins were up-regulated, and 20 proteins down-regulated under all conditions (Figure 5, Table 4). The most highly up-regulated DEPs shared between the four conditions include 4-aminobutyrate aminotransferase (SC04_2868; fold change 1.81-2.33), transaconitate 2-methyltransferase (SC04_3021; fold change 1.59-2.12), and glucosamine-fructose-6-phosphate aminotransferase (GlmS) (SC04_0181; fold

change 1.54-1.78). 4-aminobutyrate aminotransferase catalyses transamination between primary amines and α -keto acids; trans-aconitate 2-methyltransferase metabolises trans-aconitate, a toxic metabolite derived from the TCA cycle (Cai *et al.*, 2001), GlmS catalyses the conversion of fructose-6-phosphate to glucosamine-6-phosphate (GlcN-6P). GlcN-6P is a central metabolite that stands at the crossroads of aminosugar metabolism, glycolysis, nitrogen metabolism and cell-wall synthesis in *Streptomyces* (Urem *et al.*, 2016). For GlcNAc and Chitosan treated samples, increased GlcN-6P levels were expected, but this does not explain the up-regulation of GlmS. Among the most strongly down-regulated proteins common to all four conditions, there were two MarR-type regulatory proteins, SC04_2720 (fold change -3.72 to -2.70) and SC05_0108 (foldchange -1.22 to -1.90). MarR family regulators typically control stress-related regulons, including resistance to multiple antibiotics, organic solvents and oxidative stress agents (Alekhshun *et al.*, 2001, Alekhshun & Levy, 1999). Although the effects of the clustered small molecules were not the same, the proteome level similarities still provide valuable information for further studies.

4

Table 2. Number of DEPs in samples treated with different small molecules.

Small molecules	Abbreviation	Up-regulated	Down-regulated	Total DEPs
Hydroxycoumaric acid	HyC	67	92	159
Benzoic acid	BA	44	39	83
Chitosan	CS	36	36	72
Jasmonic acid	JA	35	28	63
N-acetylglucosamine	GlcNAc	25	30	55
Indole-3-acetic acid	IAA	23	18	41
Naphthaleneacetic acid	NAA	18	15	33
Butyrate	BR	8	15	23
Salicylic acid	SA	12	4	16
Cellobiose	CB	7	2	9
Ferulic acid	FA	2	1	3
Cinnamic acid	CA	0	0	0

Table 3. Overrepresented gene ontology (GO) biological process terms in hydroxycoumarin treated samples.

DEP sets	GO ID	Class name	<i>p</i> -value	Test set	Reference set	Proteins in test set
Up	0015986	ATP synthesis coupled proton transport	6.48×10^{-3}	3/31	5/891	SC04_0848 SC04_2604 SC04_2607
All	0015986	ATP synthesis coupled proton transport	3.79×10^{-2}	3/61	5/891	SC04_0848 SC04_2604 SC04_2607

Table 4. Common DEPs for the PCA cluster B treatments

ID	Log ₂ fold change					Annotation
	BA	CS	GlcNAc	JA	Mean	
sc04_2868	2.16	2.33	1.81	2.13	2.11	4-aminobutyrate aminotransferase
sc04_3021	2.12	2.12	1.59	2.07	1.98	Trans-aconitate 2-methyltransferase
sc04_2231	1.79	1.84	1.98	1.61	1.8	Phosphoenolpyruvate carboxykinase [GTP]
sc04_3260	1.93	1.82	1.7	1.65	1.77	50S ribosomal protein L16
sc04_0181	1.67	1.63	1.78	1.54	1.66	L-glutamine-D-fructose-6-phosphate amidotransferase
sc03_0574	1.41	1.33	1.26	1.26	1.31	MMPL domain protein
sc04_2879	1.44	1.17	1.29	1.07	1.24	Proline-tRNA ligase
sc02_0728	1.31	1.27	1.28	1.11	1.24	Alkaline d-peptidase
sc04_2093	1.21	1.1	1.36	1.26	1.23	Uncharacterized protein
sc07_0017	1.2	1.14	1.18	1.25	1.19	Delta-aminolaevulinic acid dehydratase
sc04_0048	1.16	1.05	1.16	1.01	1.1	DNA-binding protein HU 1
sc05_0086	1.04	1.12	1.07	1.14	1.09	Zinc protease
sc03_2044	1.17	1.06	1.08	1.05	1.09	Uncharacterized protein
sc04_1365	-1.24	-1.11	-1.19	-1.03	-1.14	Soj family protein
sc04_0799	-1.18	-1.17	-1.16	-1.09	-1.15	DUF88 domain-containing protein
sc04_0177	-1.2	-1.11	-1.21	-1.12	-1.16	Oligoribonuclease
sc04_1335	-1.33	-1.21	-1.3	-1.12	-1.24	Carbamoyl-phosphate synthase large chain
sc06_0096	-1.14	-1.36	-1.15	-1.43	-1.27	Non-specific serine/threonine protein kinase
sc03_0077	-1.6	-1.41	-1.13	-1.2	-1.33	Putative serine/threonine protein kinase
sc04_2018	-1.53	-1.25	-1.4	-1.33	-1.38	Enoyl-CoA hydratase/isomerase
sc05_0108	-1.22	-1.23	-1.23	-1.9	-1.39	MarR-family transcriptional regulator
sc07_0025	-1.57	-1.55	-1.24	-1.38	-1.44	Putative oxidoreductase
sc06_0206	-1.7	-1.45	-1.5	-1.45	-1.53	ATP-dependent chaperone ClpB
sc04_2325	-1.53	-1.54	-1.55	-1.58	-1.55	Oligopeptide/dipeptide ABC transporter ATPase subunit
sc03_1237	-1.84	-1.55	-1.49	-1.42	-1.58	LclR family transcriptional regulator
sc03_1622	-1.42	-1.21	-2.01	-1.69	-1.58	Cold-shock DNA-binding domain protein
sc04_1597	-1.91	-1.81	-1.64	-1.52	-1.72	Putative cysteinyl-tRNA synthetase
sc08_0001	-1.67	-1.79	-1.92	-1.56	-1.73	Phosphoenolpyruvate-dependent sugar phosphotransferase
sc04_0830	-1.8	-1.77	-1.69	-1.7	-1.74	Glutamate synthase subunit beta
sc06_0083	-1.84	-1.98	-1.71	-1.85	-1.85	Phenylacetic acid degradation protein PaaN
sc04_2933	-2.04	-1.99	-1.7	-1.82	-1.89	Acyl-CoA synthetase AMP-forming /AMP-acid ligase II-like protein
sc07_0022	-2.12	-1.92	-2.12	-2.02	-2.04	Putative secreted serine-rich protein
sc04_2720	-3.72	-3.44	-3.71	-2.7	-3.39	Regulatory protein MarR

***S. roseifaciens* efficiently converts JA to JA-Gln in liquid cultures**

It was shown previously that plant hormone JA is able to elicit antibiotic production in *S. roseifaciens*, and that *S. roseifaciens* aminoacylates JA to form jasmonoyl-glutamine (JA-Gln) (jasmonoyl-glutamine; van der Meij, 2020). To investigate the effect of JA on *S. roseifaciens*, and to identify possible enzymes involved in the conversion of JA to JA-Gln, an induction experiment was carried out. A 500 mL pre-culture was set up, and sampling started after 16 h ($t=0$, early exponential phase). After three samples were taken, the culture was evenly divided over 30 smaller culture-flasks with or without JA (15 flasks each). For both experimental and control group, mycelia of three cultures were harvested after 15 min, 30 min, 1 h, 2 h, and 4 h. For each sample, a small part was kept for proteomics, the rest was extracted for metabolomic analysis, to investigate the conversion of JA-Gln.

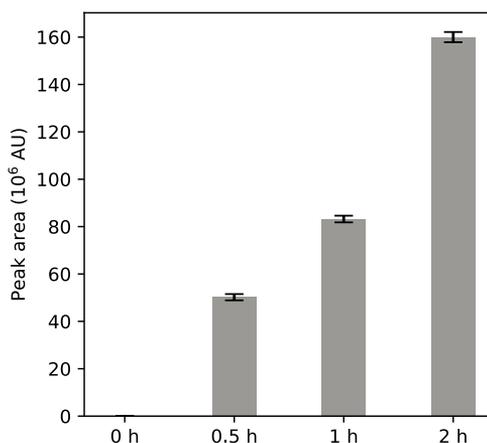


Figure 6. Relative intensity of jasmonoyl-glutamine at different time points.

Mycelium was grown in TSBS, a pre-culture was grown for 16 h before JA treatment ($t = 0$ h). The time after JA treatment was shown on the x axis. Samples were taken in parallel with proteomics samples. The same amount of extract was injected into the LC-MS system. Error bars indicate the standard errors of three replicates.

The formation of JA-Gln was previously identified by both MS and NMR in the extracts of *S. roseifaciens* grown in the presence of JA for 5 days (van der Meij, 2020). JA-Gln was detected using LC-MS by its retention time, together with its MS and MS² spectra. In the MS spectra, the $[M + H]^+$ ion of the compound was observed at m/z 339.1925. As for the MS² spectra, the most characteristic fragments observed were the $[M + H]^+$ ions of Gln (m/z 147.08) and JA-H₂O (m/z 193.12) (Figure S1), confirming this compound is indeed JA-Gln. The peak area of JA-Gln was extracted for semi-quantification. The results show that JA-Gln was

clearly detected 30 min after induction, increasing 1- and 2-fold at 1 h and 2 h after induction, respectively (Figure 6). Unfortunately, JA does not ionize well in the positive mode, so it was not detected in the LC-MS analysis performed.

Effect of JA on the proteome of *S. roseifaciens*

Pearson's correlation coefficients were calculated for proteomics data and a cut-off score of 0.97 was maintained for the correlation within triplicates. As a result, control sample at 1 h after induction was removed. The data normalization and quantification were done again using the rest of samples, ensuring best data quality. Subsequently, single quantifications amongst replicates were removed. In total, 2,137 proteins were quantified in at least one replicate. For each sample, the number of quantified proteins ranged from 1,981 to 2,058. New correlation coefficients were calculated between samples (Figure S2). The 4 h samples stood out the most in comparison to the other samples. No clear difference was observed from the correlations between JA and control samples indicating a clear shift in the proteome of *S. roseifaciens* after 2 h. Comparing JA treated samples with control samples at the same time points, only 17 DEPs conformed to the criterion of two-fold difference in protein levels and a p -value < 0.05 in T-test (Figure 7, Table S2).

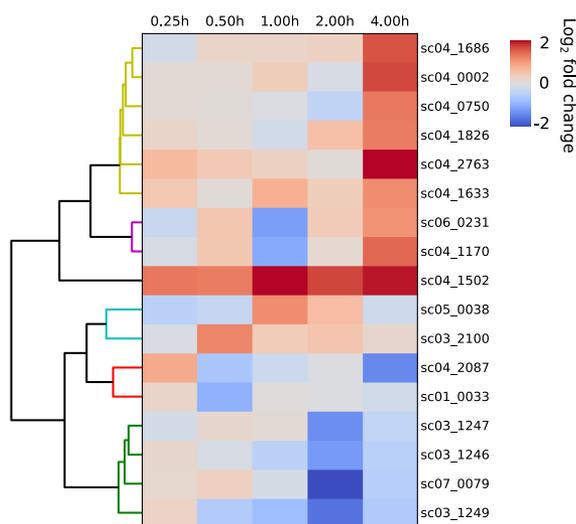


Figure 7. Heatmap showing \log_2 fold change of protein level comparing samples from JA-treated cultures and control cultures. Only proteins significantly changed (\log_2 ratio ≥ 1 , p -value < 0.1) are shown and clustered based on the hierarchical distance of the profile.

Jar enzymes are suppressed by JA

Among the 17 DEPs, SC03_1246, SC03_1247, SC03_1249, and SC07_0079 formed a cluster that were down-regulated 2 h after induction in the JA-treated

4

samples (Figure 7). Proteins numbered from SC03_1245 to SC03_1249 all showed a similar pattern (Figure 8A), suggesting they are encoded by a single gene cluster. *In-silico* operon analysis suggested that the genes SC03_1242 to SC03_1253 most likely form a single operon, which we designated as the JA-responsive (*jar*) cluster (Figure 8B). SC03_1245 to SC03_1249 were nominated JarA-E. Interestingly, JarA, JarB and JarC share high homology with the paerucumarin biosynthetic enzymes PvcA, PvcB, and PvcC (BLAST positives 63%, 64%, 78%; *e*-value 9×10^{-100} , 3×10^{-96} , 0, respectively) from *Pseudomonas aeruginosa* (Clarke-Pearson & Brady, 2008). Paerucumarin is an isocyanide functionalized coumarin that affects biofilm development by *Pseudomonas aeruginosa*, possibly through modulating the intracellular signalling pathways including the extracellular ion binding pathway (Qaisar *et al.*, 2016). The biosynthesis of paerucumarin requires PvcABCD, with PvcC and PvcD functioning as two-component flavin adenine dinucleotide-dependent monooxygenases (Clarke-Pearson & Brady, 2008). In *S. roseifaciens*, an orthologue of *pvcD* is missing. In *E. coli*, however, the orthologue of PvcC/JarC suffices for oxidation of its substrate. We therefore propose that JarABC may catalyse the formation of a paerucumarin-like compound, whereby the biosynthetic enzymes are suppressed in response to JA. The *jar* cluster also encodes a putative OmpR/PhoB type transcriptional regulator (SC03_1244), similar to the global regulator of antibiotics AfsR in *S. coelicolor* (Blast identity 34%). Like AfsR, SC03_1244 has a DNA binding domain and a BTAD domain which possibly mediates its interaction with other regulators (Yeats *et al.*, 2003, Floriano & Bibb, 1996, Horinouchi, 2003). It is yet unclear if this regulator controls the expression of the *jar* cluster.

To analyse the possible role of the *jar* cluster in the response to JA, a *jarA* knock-out mutant was created using CRISPR-Cas9 by removing the +65 to +913 region relative to the translational start of *jarA*. Agar diffusion assays with *E. coli* ASD19 as the indicator strain revealed that *jarA* mutants had lost the ability for its antimicrobial activity to be induced by JA (Figure 9). This suggests that indeed the *jar* operon plays a role in this process. One hypothesis is that the metabolite produced by the Jar proteins may regulate antibiotic production by *S. roseifaciens* in a concentration-dependent manner, with an inverse correlation between the metabolite concentration and antibiotic activity. Further work is needed to identify the nature of the metabolite produced from the *jar* gene cluster and elucidate the signal transduction pathway from JA sensing to the elicitation of antibiotic production.

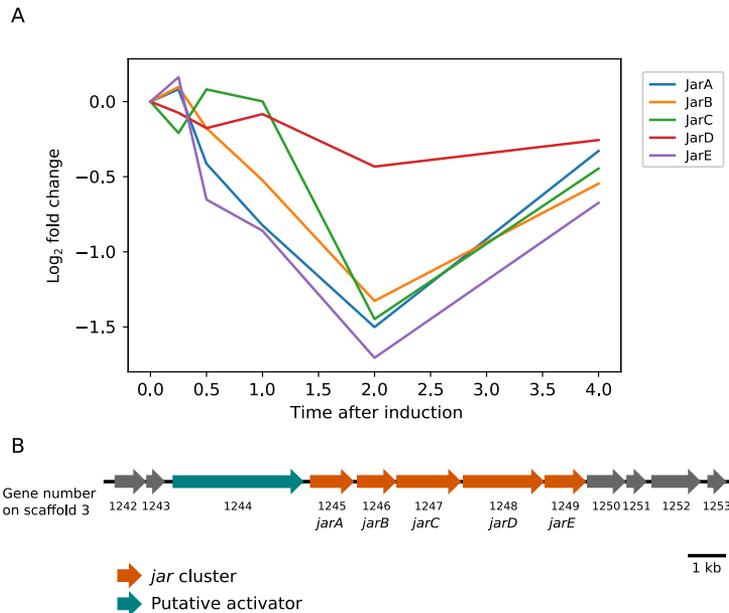


Figure 8. Jar proteins suppressed by JA and gene organization of the jar gene cluster in *S. roseifaciens*. A) Log₂ fold change of the average protein levels between jasmonic acid treatment and control sample. B) Genes of the *jar* cluster (coloured genes) and its surrounding genes on the genome.

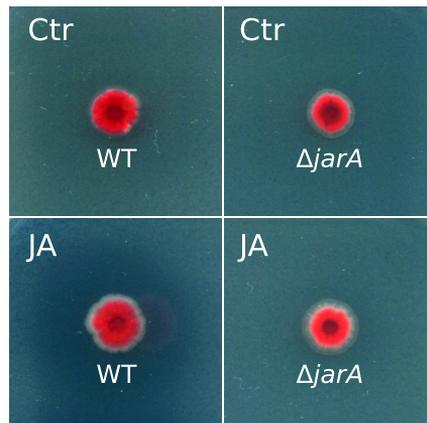


Figure 9. Antimicrobial assays for *jarA* knock-out mutant of *S. roseifaciens*. Spores of *S. roseifaciens* were spotted on MM agar media supplemented with 1% glycerol and 0.5% mannitol and grown for 5 days before overlay with the indicator strain *E. coli* ASD19. Pictures were taken 16 h after overlay.

JA can promote the level of an MFS transporter

Among the 17 DEPs, SC04_1502 was the only protein that was up-regulated at all time points from 15 min to 4 h. SC04_1502 is a putative major facilitator superfamily (MFS) transporter with 14 transmembrane domains. Functional analysis of SC04_1502 showed that this protein is a tetracycline-type resistance protein/drug resistance transporter and a likely transporter of small molecules (InterPro entry IPR001411). MFS transporters have varying substrate specificity. High specificity proteins include the tetracenomycin exporter TcmA of *Streptomyces glaucescens* (Martín *et al.*, 2005) and the riboflavin exporter RfnT in *Rhizobiaceae* (Vitreschak *et al.*, 2002); lower specificity proteins include the multiple drug resistance efflux pump EmrB in *E. coli* (Lomovskaya & Lewis, 1992) and the siderophore iron transporter Str1 in *Schizosaccharomyces pombe* (Pelletier *et al.*, 2003).

4

To test if the transporter may be involved in the transport of JA or JA-Gln, a SC04_1502 knock-out mutant was made by replacing the +372 to +1628 region (relative to the translational start site) with the apramycin resistance cassette, followed by deletion of the cassette using Cre recombinase so as to avoid polar effects. Subsequently, the in-frame deletion mutant was tested for its response to JA. While the antimicrobial activity of the mutant was decreased as compared to the parental strain, the bioactivity was still inducible by JA (Figure S3). However, this does not necessarily rule out the possibility of the involvement of SC04_1502 in transporting JA and/or JA-Gln. For example, its function may be compensated by other enzymes such as the highly abundant SC04_3061 (Figure S4) or by diffusion. Further work is needed to elucidate the function of the MFS family transporter SC04_1502 and its role in the response of JA.

Conclusion

In this chapter, the response of *S. roseifaciens* to different candidate elicitors of natural product biosynthesis was compared using quantitative proteomics. Selected candidates include plant hormones and sugars derived from natural polysaccharides. Of these, hydroxycoumarin induced a response that is distinct from that of the 12 other growth conditions. This may be caused by changes in energy flow indicated by the changes in ATP generating enzymes. Jasmonic acid, N-acetylglucosamine, chitosan and benzoic acid induced a surprisingly similar response, the basis for which remains unknown. Comparison of the proteomic changes revealed the up-regulation of a few metabolic enzymes. GlmS, which connects glycolysis to GlcNAc metabolism was among the most strongly up-regulated enzymes in all samples within this group. The proteomics data also show two MarR-family regulators that were down-regulated.

We paid special attention to the plant hormone JA, as previous data obtained in our laboratory showed that many Actinobacteria respond to JA by enhancing their antibiotic production (Anne van der Meij, PhD thesis). In this research, we found one BGC that was down-regulated in *S. roseifaciens* when JA was added and was therefore designated *jar* to highlight its responsiveness to JA. The *jar* cluster shares homology with the BGC for paerucumarin production in *P. aeruginosa*. A *jarA* mutant was made, which indeed showed altered antibiotic production and hence its possible involvement in the JA response. In fact, *S. roseifaciens* had lost its ability to respond to JA, whereby the induction of antibiotic production by JA had been lost in the *jarA* null mutant. A second candidate element of the JA signalling pathway that we identified was an MFS-family transporter, whose level was significantly promoted by JA. The precise roles of these proteins in JA sensing and the resulting change in antibiotic production awaits further investigation.

Material and methods

Strains and culturing conditions

S. roseifaciens, previously *Streptomyces* sp. MBT76 (van der Aart *et al.*, 2019), was used in this study. *E. coli* ASD19 (Avalos *et al.*, 2018) was used as an indicator strain in antimicrobial activity assays. *E. coli* JM109 was used for routine cloning. Methylation deficient strain *E. coli* ET12567 containing driver plasmid pUZ8002 (Paget *et al.*, 1999) was used in conjugation experiments for introducing DNA into *Streptomyces*.

For solid-grown cultures, spores collected from 14 days MYM culture of *S. roseifaciens* were used. MM agar (Kieser *et al.*, 2000) supplemented with 0.5% mannitol and 1% glycerol was used for proteomics culture preparation. Approximately 10^7 spores were inoculated on each plate covered with cellophane and grown at 30°C for 48 h. Then the cellophane disks were carefully moved to a new plate containing desired concentration of small molecules (Table S5) and grown for another 48 h. Mycelium was then collected using a small spatula from the cellophane and transferred to a 2 mL EP tube with two 4 mm diameter ion beads. The samples were snap-frozen in liquid nitrogen and stored in -80°C for protein extraction.

S. roseifaciens mycelial stocks were made by inoculating 500 μ L initiator mycelium in 250 mL TSBS medium in a 2 L Erlenmeyer flask equipped with a steel spring, this culture was grown at 30°C with constant shaking at 200 rpm. After 2 days of growth, mycelium was collected by centrifugation and resuspended in 20% glycerol and aliquots was made and stored at -80°C. 1.5 mL *S. roseifaciens* mycelium stock was inoculated into 500 mL pre-warmed TSBS medium in a 3 L Erlenmeyer flask equipped with a spring and culture in a 30°C shaking incubator with a constant shaking speed of 200 rpm. After 16 h growth, 3 \times 15 mL culture was collected as sample time 0, the rest of cultures was divided per 15 mL to new 100 mL flasks with JA or with equal amount of solvent (ethanol). The final concentration of JA was 0.01% (w/v). At each of the following time point: 15 min, 30 min, 1 h, 2 h and 4 h, samples were collected in triplicate. For each sample, 2 mL was pelleted for proteomics sample preparation, the remaining volume was used for metabolomics analysis.

Knockout mutants of *S. roseifaciens*

Oligonucleotides used in this study was listed in Table S3. PCR was preformed using Pfu DNA polymerase using standard protocol (Colson *et al.*, 2007). All plasmid and constructs used for generating knockout mutants are summarized in Table S4. The constructs generated in this study were verified by Sanger sequencing performed in BaseClear (Leiden, The Netherlands). *E. coli* strains were grown in Luria broth at 37°C supplemented with the appropriate antibiotics

(ampicillin, apramycin, kanamycin and/or chloramphenicol at 100, 50, 25 and 25 $\mu\text{g}\cdot\text{mL}^{-1}$, respectively) depending on the vector used (Table S4).

Knockout of *jarA* was performed using pCRISPR-Cas9 technology (Tong *et al.*, 2015a). A spacer sequence was designed at the middle of *jarA*, with the sequence of ACAGGTCCTCCAGCATGAAG. The design of sgRNA guides Cas9 to make a double strand brake at around nt position +610 relative to the *jarA* translational start site. After spacer integration, the PCR product of *jarA* flanking regions for homology-directed repair designed to knock out the +65 to +913 regions (relative to the *jarA* translational start site) were inserted and a construct made using Gibson assembly. The resulting knockout construct was designated pGWS1460. This construct was introduced into *S. roseifaciens* through conjugation. Loss of the pGWS1460 was confirmed by the apramycin sensitivity of the exconjugants, knockout of *jarA* was further confirmed by PCR.

Knockout of the MFS family protein SC04.1502 was performed using pWHM3-oriT system which was a conjugatable version of pWHM3 (Vara *et al.*, 1989). The knockout strategy was essentially the same as described previously (Zhang *et al.*, 2018), conjugation was used to introduce construct to *S. roseifaciens*. Basically, around 1500 bp of upstream and downstream region of SC04.1502 was amplified by PCR from *S. roseifaciens* genome. The upstream region was thereby cloned as *HindIII-XbaI* fragment, and the downstream region was cloned as *XbaI-EcoRI* fragment. The apramycin resistance cassette *aac(3)IV* with flanking *loxP* sites was digested from construct made in previous study (Zhang *et al.*, 2018). These three fragments were ligated into *EcoRI-HindIII*-digested pWHM3-oriT. The presence of the *loxP* recognition sites allows the efficient removal of the apramycin resistance cassette following the introduction of a plasmid pUWL-Cre expressing the Cre recombinase (Fedoryshyn *et al.*, 2008, Khodakaramian *et al.*, 2006). Knock-out construct pGWS1461 was created for the deletion of nucleotide positions +372 to +1628 of SC04_1502, where +1 refers to the translation start site of SC04_1502. Introducing this construct to *S. roseifaciens* followed by losing this construct resulted in replacing the respective region with apramycin resistance cassette, which was subsequently removed using Cre expressing construct pUWLCre. The confirmation of knockout and losing the apramycin resistance cassette was confirmed by PCR followed by Sanger sequencing of the product.

Metabolomics analysis

For each sample collected from liquid grown cultures, about 0.5 g of Diaion[®] HP20 (Resindion, Binasco, Italy) was added and the tube was then placed on a rotator at 4°C overnight. The HP20 beads were then collected on glass wool, washed with water, and then soaked in methanol overnight at room temperature. The methanol was evaporated under vacuum at 40°C and the residue was dissolved in methanol again at a fixed concentration of 1 $\text{mg}\cdot\text{mL}^{-1}$ for analysis.

LC-DAD-HRESIMS spectra were obtained using a Waters Acquity UPLC system, equipped with Waters Acquity PDA, and coupled to a Thermo Instruments MS system (LTQ Orbitrap XL). The UPLC system was run using Acquity UPLC HSS T3 C₁₈ column (1.8 μm, 100 Å, 2.1 × 100 mm). Solvent A was 0.1% formic acid, 95% H₂O and 5% acetonitrile. Solvent B was 0.1% formic acid, 95% ACN and 5% H₂O. The gradient used was 2% B for 0.5 min, 2-40% for 5.5 min, 40-100% for 2 min, and 100% for 3 min. The flow rate was 0.5 mL·min⁻¹. The MS conditions used were capillary voltage 5 V, capillary temperature 300 °C, auxiliary gas flow rate 5 arbitrary units, sheath gas flow rate 50 arbitrary units, spray voltage 3.5 kV, mass range 100-2000 *m/z*, FT resolution 30000. The MS spectra were acquired in the positive mode.

Sample preparation for proteomics

4 For mycelium collected from plates, the 2 mL EP tubes containing mycelium and metal beads were put into an adaptor and disrupted using Tissuelyser II (Qiagen, Venlo, The Netherlands) for 2 cycles of 30 s at 30 Hz. The samples were kept frozen during this process by cooling using liquid nitrogen before and in between cycles. 0.6 mL disrupting buffer (4% SDS, 0.06 M DTT, 100 mM Tris-HCl pH 7.6, 50 mM EDTA) is then added. Samples were vortexed to allow mixing with disrupting buffer and restore to room temperature. Cell debris was then removed by centrifugation at 16,000 × *g* for 10 minutes at 20°C, 50 μL supernatant was taken for protein precipitation.

For liquid grown cultures, samples were centrifuged at 5,000 × *g* for 10 min, cell pellets were resuspended in 1 mL disrupting buffer and transferred to 1.5 mL screw-cap tube with about 0.5 g 0.1 mm glass beads. All samples were disrupted using FastPrep FP120 (Thermo, Pennsylvania, U.S.) at speed 6.0 (m·s⁻¹) for 5 times 30 s disruption in a 4°C cold room, with 5 minutes in between runs to cool down the samples. Samples after disruption were centrifuged at 16,000 × *g* for 10 minutes at 20°C to remove cell debris. 50 μL supernatant was taken for protein precipitation.

Protein extracts from both liquid and plate grown cultures were precipitated using chloroform-methanol method described by (Wessel & Flugge, 1984). Briefly, 200 μL methanol was added to 50 μL sample and mixed, then 100 μL chloroform was added and mixed. 150 μL water was then added and vortexed for 30 s. Phase separation was achieved by centrifugation at 16,000 *g* for 5 minutes. After removing the water phase, 150 μL methanol was added to the remainder followed by vortex and centrifugation at 16,000 *g* for 5 minutes to precipitate and pellet proteins. Supernatant was then removed, and the protein pellets were then dried in a vacuum centrifuge.

Protein pellets were resuspended in 50 μL of 0.05 % RapiGest SF solution (in 50 mM ammonium bicarbonate) and incubated at 95°C for 5 minutes to dissolve as

much protein as possible. 14 μL was taken for BCA assay to determine protein concentration. 5 mM DTT was added in the rest of samples and incubate at 60°C for 30 minutes. Iodoacetamide was then added to 21.6 mM, and samples were incubated at room temperature in dark. Trypsin was then added at an enzyme to substrate ratio of 1:100 allowing overnight digestion at 37°C. Samples were then acidified to degrade RapiGest by adding 5% TFA to 0.5% and incubated at 37°C for 60 minutes, followed by centrifugation at 16,000 g for 10 min. The supernatant was desalted using STAGE-tip microcolumns as described previously (Gubbens *et al.*, 2014, Rappsilber *et al.*, 2007). Briefly, 8 μg of acidified peptides of each sample was loaded on STAGE-tip, washed twice with 0.5% formic acid solution, eluted with elution solution (80% acetonitrile, 0.5% formic acid). Acetonitrile was then evaporated in a SpeedVac (Thermo, Pennsylvania, U.S.). Final peptide concentration was adjusted to 100 $\text{ng}\cdot\mu\text{L}^{-1}$ using sample solution (3% acetonitrile, 0.5% formic acid) before analysis.

UPLC and mass spectrometry measurement and data processing

200 ng (2 μL) digested peptide was injected and analysed by reversed-phase liquid chromatography on a nanoAcquity UPLC system (Waters, Massachusetts, U.S.) equipped with HSS-T3 C18 1.8 μm , 75 μm X 250 mm column (Waters, Massachusetts, U.S.). A gradient from 1% to 40% acetonitrile in 110 min (ending with a brief regeneration step to 90% for 3 min) was applied. [Glu¹]-fibrinopeptide B was used as lock mass compound and sampled every 30 s. Online MS/MS analysis was done using Synapt G2-Si HDMS mass spectrometer (Waters, Massachusetts, U.S.) with an UDMS^E method set up as described in (Distler *et al.*, 2014)

Raw data from all samples were first analysed using the vendor software ProteinLynx Global SERVER (PLGS) version 3.0.3. Generally, mass spectrum data were generated using an MS^E processing parameter with charge 2 lock mass 785.8426, and default energy thresholds. For protein identification, default workflow parameters except an additional acetyl in N-terminal variable modification were used. Reference protein database was made from predicted 7974 coding sequences described in (Wu *et al.*, 2016c). The resulted dataset was imported to ISOQuant (version 1.8, Distler *et al.*, 2014) for label-free quantification. ISOQuant processing includes spectrum alignment, data normalization at spectrum level, peptide quantification by top3 method and subsequent protein quantification. Default high identification parameters were used in the quantification process.

Acknowledgments

We would like to acknowledge Bogdan I. Florea of Leiden University, Leiden, The Netherlands, for running and for monitoring the proteome measurements, and the

bio-organic synthesis group at Leiden University for providing the opportunity to use their instrumentation.