Signalling pathways that control development and antibiotic production in streptomyces

Urem, M.; Urem M.

CHAPTER VI

THE ROK-FAMILY REGULATOR ROKL6 (SCO1447) IS INVOLVED IN THE CONTROL OF GLUCOSAMINE-SPECIFIC METABOLISM IN *S. COELICOLOR*

Mia Urem, Magdalena A. Świątek-Połatyńska, Eva J. van Rooden, Bobby I. Florea and Gilles P. van Wezel

ABSTRACT

The aminosugar *N*-acetylglucosamine (GlcNAc) is an important pleiotropic signalling molecule for the regulation of development and antibiotic production in *Streptomyces coelicolor*. Our data suggest that the deacetylated aminosugar glucosamine (GlcN) is metabolised via the GlcNAc pathway and that the ROK-family protein RokL6 (SCO1447) is a GlcN-specific regulator. Deletion of *rokL6* relieved the toxicity of GlcN to *nagB* mutants, while the mutant still failed to grow on GlcNAc. Initial proteomic comparison of *S. coelicolor* M145 and its *rokL6* null mutant revealed strongly enhanced expression of sporulation regulator WhiB, while Red biosynthetic proteins were down-regulated. This suggests that RokL6 directly or indirectly controls expression of genes involved in development and antibiotic production. A possible autoregulatory RokL6 binding site was discovered in the intergenic region between *rokL6* and SCO1448. The latter encodes an MFS-transporter with unknown substrate, but a direct role in GlcN transport could not be ascertained. Surprisingly, the *nagB-rokL6* double mutant was resistant to the toxic glucose analogue 2-deoxyglucose (DOG), which is a glycolytic inhibitor and promising chemotherapeutic drug. Our results expose intricate links between GlcN, GlcNAc and glucose metabolism, thereby shedding new light on the regulation and pathways for (amino-)sugar metabolism in streptomycetes.

Introduction

Secondary metabolite production in the filamentous streptomycetes is closely linked to the switch from vegetative to aerial growth (Bibb, 2005; van Wezel & McDowall, 2011). At the onset of development, the vegetative mycelium is partially degraded via programmed cell death (PCD) so as to provide the building blocks necessary for aerial growth in an otherwise nutrient-depleted environment (Manteca *et al.*, 2005; Miguelez *et al.*, 1999; Park & Uehara, 2008). During cell-wall recycling, the subunits of the peptidoglycan (PG), namely *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) that make up the PG strands and the cross-linking amino acids, are re-imported into the cell. GlcNAc and its metabolism plays a key role in signalling during the control of development and antibiotic production in streptomycetes, as extensively described in Chapter II.

Streptomyces are saprophytic organisms that degrade many different polysaccharides, including cellulose, chitin and chitosan, which releases glucose (Glc), GlcNAc and glucosamine (GlcN), respectively. While metabolic pathways of glucose and GlcNAc have been well studied (Świątek *et al.*, 2012a; Nothaft *et al.*, 2010; Angell *et al.*, 1994; Gubbens *et al.*, 2012; van Wezel *et al.*, 2007; van Wezel *et al.*, 2005; Colson *et al.*, 2008; Xiao *et al.*, 2002), little is known of how GlcN is metabolised by streptomycetes (Chapter V). The hydrolysis of chitosan and the metabolism of chitosan-derived oligomers [$(GlcN)_{2-3}$] is controlled by CsnR (Dubeau *et al.*, 2011; Viens *et al.*, 2015), a repressor from the ROK (Repressors, ORFs and Kinases) family of transcriptional regulators (Titgemeyer *et al.*, 1994). $(GlcN)_{2-3}$ oligomers are imported via the ABC-transport complex CsnEFH-MsiK and are then presumably hydrolysed and phosphorylated by CsnH, a sugar hydrolase, and CsnK, a ROK-family kinase, respectively. However, the transporter of monomeric GlcN has not yet been identified.

In the absence of the NagB enzyme, high concentrations of GlcN or GlcNAc are toxic to *Streptomyces coelicolor* (Świątek *et al.*, 2012b; Świątek, 2012). This suggests that streptomycetes are not able to cope with high concentrations of GlcN-6P, which accumulates when NagB activity is low. Taking advantage of this system, a screen of GlcN-derived suppressor mutants was performed, resulting in the identification of genes related to GlcN metabolism (Świątek *et al.*, 2012b; Świątek, 2012; Chapter V). An unexpected discovery was that several suppressor mutations were found in the presumably GlcNAc-specific metabolic genes *nagA* and *nagK,* for GlcNAc deacetylase NagA and GlcNAc kinase NagK, respectively (Świątek *et al.*, 2012a; Chapter V). Since both genes are specific for GlcNAc, these data suggest a unique metabolic pathway for GlcN in *S. coelicolor*. GlcN is thereby converted to GlcNAc and then further metabolised via the GlcNAc pathway. This suggests the presence of a yet unidentified GlcN *N*-acetylase in *Streptomyces*. Interestingly, several independent suppressor mutations were found in the gene for the ROK-family regulator RokL6 (SCO1447). Deletion of *rokL6* specifically restored growth to *nagB* mutants in the presence of GlcN but not GlcNAc. The gene SCO1448 for a major facilitator superfamily (MFS) transporter, which is located divergently from *rokL6* on the genome, was subsequently also deleted in the *nagB* null mutant. However, this did not restore growth on either aminosugar (Świątek, 2012; Chapter V).

In this work, we investigated the role of GlcNAc-metabolic enzymes and RokL6 in GlcN metabolism. The regulatory role of RokL6 was investigated by comparing the proteomes of wild-type *S. coelicolor* and its *rokL6* deletion mutant in the absence and presence of GlcN.

Materials and Methods

Bacterial strains, culture conditions, plasmids and oligonucleotides.

The bacterial strains used in this work are listed in Table S1 of the supplemental material. *Escherichia coli* was transformed according to standard procedures (Sambrook *et al.*, 1989), whereby *E. coli* JM109 was used for routine cloning and *E. coli* ET12567 (MacNeil *et al.*, 1992) to generate non-methylated plasmid DNA. *E. coli* was grown in Luria-Bertani (LB) media at 30°C and selective antibiotics were added where required to the following final concentrations; ampicillin (100 μg/ml) and chloramphenicol (25 μg/ml).

Streptomyces coelicolor A3(2) M145 was obtained from the John Innes Centre strain collection. All *Streptomyces* media and routine techniques are described in the *Streptomyces* manual (Kieser *et al.*, 2000). To cultivate mycelia for protoplast preparation, a 1:1 mixture of yeast-extract malt extract (YEME) and tryptic soy broth (TSB) was used, and protoplasts where regenerated on R5 agar. Spores were cultivated on SFM (soy flour mannitol) agar. Phenotypic characterization was done on R5 and minimal media (MM) agar supplemented with sugars as stated. Induction experiments were performed in liquid cultures with mannitol as a non-repressive carbon source; NMM (without PEG) was used for proteomics experiments. All *Streptomyces* cultures were grown at 37°C and liquid cultures were grown in flasks containing spring coils and vigorously shaken. Where required, apramycin (50 μg/ml final concentration) and/or thiostrepton (20 μg/ml final concentration) were added as selectable markers for plasmids.

All plasmids and oligonucleotides are described in Table S1 and Table S2, respectively. *S. coelicolor* M145 genomic DNA was used as a template for PCR amplification and after cloning, PCR product inserts were checked by DNA sequencing. pWHM3, an unstable multi-copy shuttle vector, was used for gene replacement strategies as described (van Wezel *et al.*, 2005), and plasmid pUWLcre, which expresses Cre recombinase, was introduced into *Streptomyces* for genomic excision via *loxP* marked sites (see below for details; Fedoryshyn *et al.*, 2008). DNA sequencing was performed by BaseClear BV (Leiden, The Netherlands).

Creation of knock-out mutants.

Deletion mutants were created for SCO1447 (*rokL6*), SCO6110 and SCO6110-6114 in different *S. coelicolor* genetic backgrounds (Table S2). Deletion mutants for *nagB* and SCO1447 (*rokL6*) in wild-type *S. coelicolor* were described previously (Świątek, 2012). Gene replacement mutants, with the gene of interest replaced by an apramycin resistance cassette, were created via homologous recombination. The upstream (left) and downstream (right) flanking regions of the gene were cloned via engineered EcoRI/HindIII restriction sites into the highly unstable vector pWHM3. The left and right flanks were PCR-amplified from the *S. coelicolor* genome using primer pairs 6110LF-1198/6110LR+6 and 6110RF+927/6110RR+2197 for the amplification of the SCO6110 flanks, and primer pairs 6114LF-1306/6114LR+9 and 6110RF+927/6110RR+2197 for the SCO6110-6114 flanks. The *aac(3)IV* apramycin resistance cassette, flanked by loxP sites, was cloned between the gene flanks via engineered XbaI sites. This resulted in gene knock-out plasmids, pGWS953 and pGWS954, for SCO6110 and SCO6110-6114, respectively, and the previously described pGWS948 for SCO1447 (*rokL6*) knock-outs. To create SCO6110 mutants, the SCO6110 knock-out plasmid pGWS953 was transformed into *S. coelicolor* M145 and its nagB or SCO1447 (*rokL6*) deletion mutants generating M145 ΔSCO6110 and the double mutants M145 ΔnagBΔSCO6110 and M145 ΔrokL6ΔSCO6110, respectively. To create SCO6110-6114 mutants, SCO6110-6114 knock-out plasmid pGWS954 was transformed into the same strains to create single mutant M145 ΔSCO6110-6114 and double mutants M145

ΔnagBΔSCO6110-6114 and M145 ΔrokL6ΔSCO6110-6114.

In all cases, correct recombination events were checked by the appropriate antibiotics resistance and, where appropriate, confirmed by PCR (Colson et al., 2007). The apramycin resistance cassette was subsequently excised by introduction of pUWLcre, a Cre recombinase expressing plasmid, which allows for efficient removal of the cassette via the loxP recognition sites to obtain deletion mutants (Fedoryshyn et al., 2008; Khodakaramian et al., 2006). Deletion mutants were checked based on the appropriate antibiotic sensitivity (loss of apramycin resistance) and confirmed by PCR.

Protein extraction and LC-MS/MS analysis.

To analyse the proteomes of *S. coelicolor* M145 and its SCO1447 (*rokL6*) null mutant in the presence and absence of GlcN, protein extracts were prepared from liquid-grown cultures as described (Gubbens et al., 2014). Pre-cultures were grown overnight in YEME and used to inoculate 50 ml of fresh NMM media (NMMP without PEG) with 25 mM Mann to an OD_{600} of around 0.1. Following 4 hours of growth, the cultures were induced with 50 mM of Mann or GlcN. Triplicate samples of 5 mL were taken 1 h after induction. For in-solution digestion of proteins with trypsin, the precipitated proteins were first resuspended in 50 µl 50 mM ammonium bicarbonate with 0.1% (w/v) RapiGest SF Surfactant (Waters) and heated at 95°C for 5 min. Following resuspension, the protein concentrations were determined using the Bio-Rad Bradford-based assay, with BSA as a reference. DTT was added to each protein sample to a final concentration of 5 mM and incubated for 30 min at 60°C. Then iodoacetamide was added to a final concentration of 21.6 mM and incubated in the dark for 30 min. Finally, trypsin was added in a 1:100 (w/w) trypsin/protein ratio and the samples were digested overnight at 37°C. To acidify the samples, formic acid was added to a final concentration of 0.5% and incubated for 30 min at 37°C, and to remove RapiGest SF, the samples were then centrifuged (20,000 g for 10 min). For further purification of the supernatant, a total of 8 µg of protein was loaded onto Stage Tips as previously described (Rappsilber et al., 2003). Acetonitrile was removed from the samples using a vacuum concentrator. Each sample was adjusted to a final concentration of around 80 ng/µL using a solution of 3% (v/v) acetonitrile and 0.1% (v/v) formic acid.

Proteomic data analysis.

Peptides were analysed on a Waters ACQUITY UPLC system (Waters, Massachusetts, USA) hyphenated to an SYNAPT G2-Si mass spectrometer (Waters, Massachusetts, USA). PicoTip emitters (OD/ID = 360/20 µm tip ID = 10 µm), trap column (C18 100 Å, 5 µM, 180 µM x 20 mm, P/N 186006527) and analytical columns (HSS-T3 C18 1.8 µM, 75 µM x 250 mm) were obtained from Waters Corporation. The mobile phases [A: 0.1% (v/v) formic acid/H2O, B: 0.1% (v/v) formic acid/acetonitrile] were made with ULC/MS grade solvents (Biosolve, Valkenswaard, the Netherlands). The emitter tip was coupled end-to-end with the analytical column via a 15 mm long TFE Teflon tubing sleeve (OD/ID 0.3 × 1.58 mm, Supelco, St Louis, MO) and installed in a stainless steel holder mounted in a nano-source base (Idex, Northbrook, IL). Not anymore.General mass spectrometric conditions were as follows: an electrospray voltage of 2.5–3.5 kV was applied to the emitter, trap collision energy 4 V, ramp transfer collision energy 20 to 45 V. Internal mass calibration was performed with [Glu1]-Fibrinopeptide B solution of 100 fmol/µL (m/z = 785.8427) as lock mass.

For shotgun proteomics analysis, 2 µl of the samples was pressure loaded on the trap column with a 5 µl min^{-1} flow for 3 min followed by peptide separation with a gradient of 60 min 5–85% B, at a flow of 0.4 µl min^{-1}. Full MS scan (50–2000 m/z) acquired in HDMSe

mode.

Data was transferred and analysed using Progenesis QI for proteomics 3.0 software (Waters, Massachusetts, USA). Peptide identification was set to a minimum 3 fragments per peptide, 7 fragments and 2 peptide per protein, and a maximum protein size of 2000 kDa.

BIOINFORMATICS ANALYSIS.

Motifs were predicted using InterProScan (Zdobnov & Apweiler, 2001), NCBI Conserved Domain Database (Marchler-Bauer *et al.*, 2015) and Pfam 24.0 (Finn *et al.*, 2008) and protein homology searches were performed using BLASTp (Altschul *et al.*, 2005). Schematic representations of *cis*-regulatory elements were produced using Boxshade (www.ch.embnet.org/software/box_form.html). Synttax was used for gene synteny (Oberto, 2013).

Results

RokL6 (SCO1447) is a conserved ROK-family protein involved in GlcN metabolism.

Putative transcriptional regulator RokL6 is characterised by an N-terminal winged helix-turn-helix DNA-binding site, found in various families of DNA-binding proteins, and a ROK-family protein domain (Fig. S1). The ROK (Repressor, ORF and Kinase) family of proteins play an important role in the (control of) sugar utilization in bacteria (Titgemeyer *et al.*, 1994). The family includes GlcNAc kinase NagK and NagC, the repressor of the *E. coli* aminosugar uptake and metabolism operon *nag* (Plumbridge, 1991); NagC also activates transcription of the *glmUS* operon, which is responsible for the synthesis of the cell-wall precursor UDP-GlcNAc (Plumbridge, 2015). In *S. coelicolor,* several ROK-family proteins have been identified that control sugar metabolism, while several also play a key role in the control of secondary metabolism (Bekiesch *et al.*, 2016; Świątek *et al.*, 2013). Disruption of *rokL6* restored growth to *S. coelicolor nagB* mutants in the presence of GlcN, strongly suggesting that RokL6 plays a key role in (the control of) GlcN metabolism (Świątek, 2012; Chapter V).

RokL6 is found in most streptomycetes, with relatively high sequence similarity between the orthologues (67-100% similarity) (Fig. S1). There is also significant gene synteny for the genomic region surrounding *rokL6* and its orthologues (Fig. 1). In particular, a gene for a major facilitator superfamily (MFS) transporter (SCO1448) often lies divergently transcribed from *rokL6*. Bioinformatic analysis using the String algorithm (Szklarczyk *et al.*, 2015) suggested a functional correlation between *rokL6* and SCO6110, a ROK-family kinase. The gene lies in an operon with genes SCO6111-6114 for an ABC transporter complex that is predicted to transport an oligopeptide.

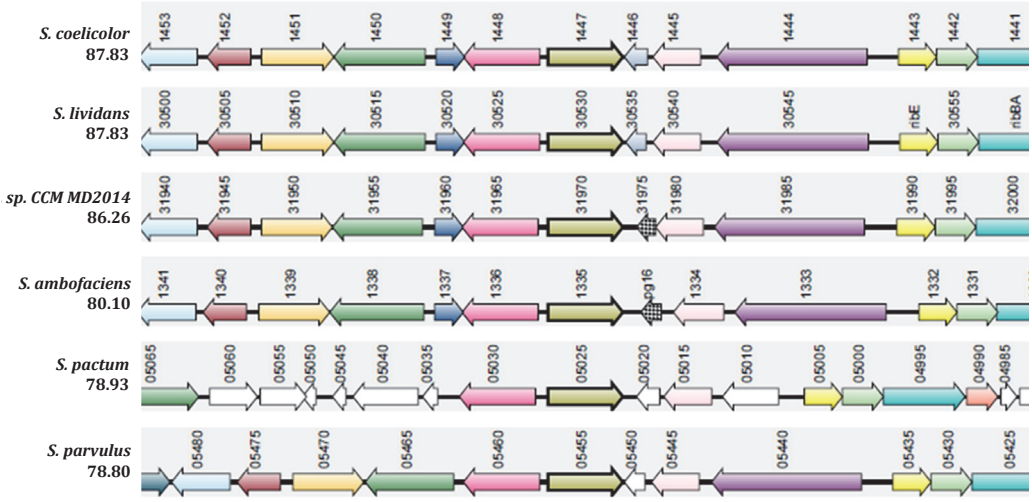


Figure 1. RokL6 conserved domains and gene synteny.
Gene synteny between RokL6 in *S. coelicolor* and homologs in other *Streptomyces* species. Synteny analysis performed by Synttax. Homologous genes are presented in the same colours.

Mutational analysis of SCO6110-6114.

Deletion mutants of SCO6110 and the entire SCO6110-6114 gene cluster were created in *S. coelicolor* by replacing the entire coding regions by the *aac(3)IV* apramycin resistance cassette via homologous recombination (see Materials and Methods for details). Mutants were selected based on apramycin resistance as a marker for the deletion and sensitivity to thiostrepton for loss of the plasmid. The resistance cassette *aac(3)IV* was flanked by *loxP*

sites, which were then used to create markerless deletion mutants via genomic excision by Cre recombinase, expressed from plasmid pUWLcre. All mutants were verified by PCR. To examine whether SCO6110 or the transporter complex SCO6111-6114 may be involved in GlcN or GlcNAc metabolism, the genes were also deleted in the *nagB* deletion mutant. To analyse possible functional linkage between SCO6110 and RokL6, a SCO6110-*rokL6* double mutant was also created by deleting SCO6110 in the *rokL6* null mutant.

Mutants with gene deletions for SCO6110 or the SCO6110-6114 operon were phenotypically characterized on MM and R5 agar with or without 50 mM GlcN or GlcNAc (Fig. 2). Deletion of neither SCO6110 nor operon SCO6110-6114 significantly altered antibiotic production or development under any of these growth conditions. In line with this, deletion of SCO6110 in the *rokL6* mutant led to a phenotype highly similar to that of the single *rokL6* mutant. The inability to restore growth of the *nagB* deletion mutation on aminosugars by deleting SCO6110 or SCO6110-6114 suggests that GlcN metabolism is not fully blocked in the mutants.

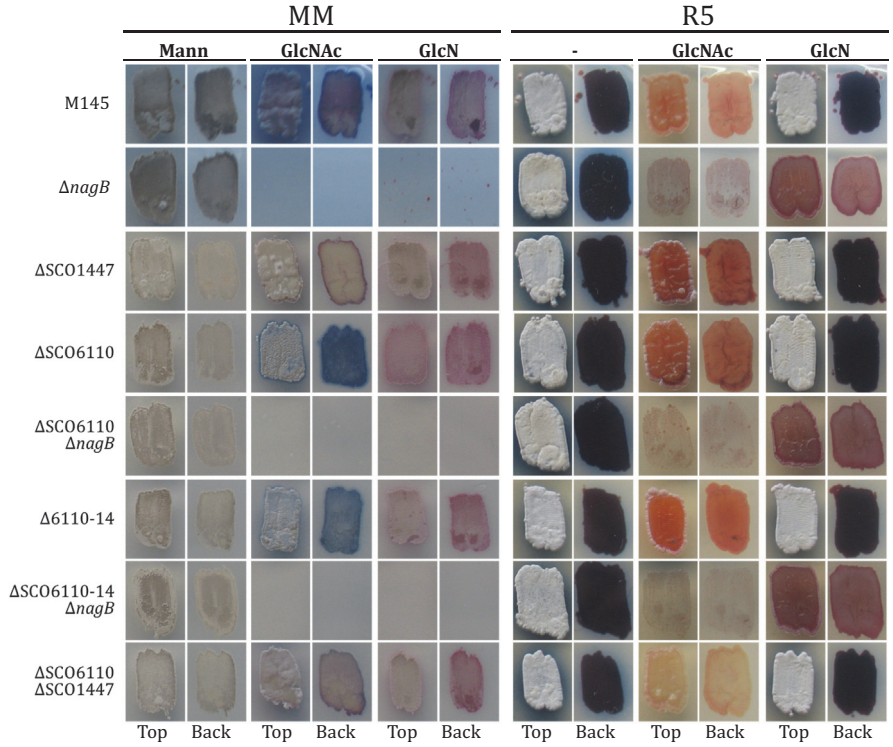


FIGURE 2. Phenotypic analysis of *S. coelicolor* mutants.
Phenotypic analysis of *S. coelicolor* wild-type (M145) and deletion mutants (Δ) as annotated on the left for each row. Spores were plated on minimal medium (MM) or rich glucose-containing media (R5) and supplemented with 50 mM mannitol (Mann), glucosamine (GlcN) or *N*-acetylglucosamine (GlcNAc) as indicated. Strains that were inhibited by GlcN or GlcNAc are indicated (red line). The top view and bottom (back) view of plates are indicated underneath.

PROTEOMIC ANALYSIS OF GLCN-SPECIFIC METABOLISM.

To assess the response of *S. coelicolor* to GlcN and potentially identify novel GlcN-metabolic enzymes, proteome analysis was done on *S. coelicolor* grown in the presence and absence of GlcN. For this, *S. coelicolor* M145 was pre-grown in YEME media, then the mycelia were washed in NMM and transferred to fresh NMM with mannitol as the sole carbon source,

to a starting OD_{600} of around 0.1. After 4 h of growth in NMM, the culture was split into six subcultures and 50 mM GlcN was added to three cultures and 50 mM Mann to the other three. A sample was collected from each subculture 1 h after induction and analysed separately to obtain three replicates. The proteomes of *S. coelicolor* grown on GlcN and Mann were then compared. The same experiment was simultaneously carried out with the *S. coelicolor rokL6* deletion mutant in order to investigate the effect of RokL6 absence on the *S. coelicolor* proteome on Mann and on GlcN, as a means to examine the potential regulatory role of RokL6.

A total of 1462 proteins were identified in all four sample groups, with an average of 1475 proteins identified per sample. This means that over 99% of all proteins were identified in all samples. Differences in protein levels between samples were only considered relevant with a fold change of 2.5 or higher and a statistically significant difference determined by ANOVA analysis using a significance level of $p < 0.1$ (Table S).

GlcN induction

Despite the large number of proteins identified, addition of GlcN significantly altered the expression of only 38 proteins in the wild-type strain, many of which of unknown function (Fig. 3A). Two well-known targets stood out, namely NagE2 (5.5 fold upregulated in the presence of GlcN) and WhiB (2.7 fold upregulated). NagE2 (SCO2907) is the membrane component Enzyme IIBC of the GlcNAc-specific PTS transporter (Nothaft *et al.*, 2010) and WhiB (SCO3034) is a regulatory protein that is required for the transition from aerial growth to sporulation (Flardh & McCormick, 2017; Willemse *et al.*, 2012; Molle *et al.*, 2000). Other categories of proteins whose expression was significantly increased in the presence of GlcN included those involved in nucleotide metabolism and members of the so-called conservon (Cvn) proteins. Additionally, protein levels of three putative transport-related proteins were significantly (around 3-fold) increased in cultures containing GlcN, namely the sugar binding protein SCO3484 and the ABC-transporter components SCO2830 (likely involved in amino acid transport) and SCO4585.

Effect of deletion of rokL6

To establish the global changes in protein expression as the result of the absence of RokL6, the proteomes of *S. coelicolor* M145 were compared to those of its *rokL6* null mutant. Again, only a small number of proteins were differentially expressed between the two strains, namely 19 proteins in the mannitol-grown cultures and 28 when GlcN was added (Fig. 3B, Table S3). In the *rokL6* mutant both RedY and RedI, which are biosynthetic proteins involved in the production of undecylprodigiosin (Red; Cerdeno *et al.*, 2001), were at least 2.5-fold decreased, suggesting that RokL6 may directly or indirectly control Red biosynthesis. Additionally, sporulation regulator WhiB again stood out, with a 4.4-fold increase in the *rokL6* mutant during growth on mannitol and 1.8-fold during growth on Mann+GlcN. Taken together, the data are consistent with induction of *whiB* by GlcN, while the gene is repressed by RokL6.

Searching for a RokL6 binding site and regulon prediction.

A bioinformatic analysis was carried out on the *S. coelicolor* genome in an attempt to identify possible DNA binding sites for RokL6. This should allow correlation of predicted RokL6 binding sites to the proteomic data. We anticipated that RokL6 auto-regulates its own gene expression and therefore the SCO1447-SCO1448 intergenic region was scrutinized for potential DNA-binding motifs. For this, the intergenic and coding regions of *S. coelicolor*

a)

GlcN-Mann		SCO#	Description
WT	Δ1447		
10.1	5.4	SCO3399	Uncharacterized protein
8.1		SCO5539	Conservon CvnB2
5.5		SCO2907	PTS component NagE2
5.4		SCO4545	Uncharacterized protein
5.1		SCO0672	Anti-sigma factor antagonist
4.6		SCO1678	Putative transcriptional regulator
4.5		SCO4610	Putative integral membrane protein
4.4		SCO4096	ATP-dependent RNA helicase
4.0		SCO0719	Uncharacterized protein
4.0		SCO5563	Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase ThiD
3.9		SCO1953	UvrABC system protein C
3.9		SCO3914	Putative transcriptional regulator
3.6		SCO3756	Putative two component system response regulator
3.5		SCO7631	Putative secreted protein
3.4		SCO6635	Bacteriophage (PhiC31) resistance PglY
3.4		SCO6169	Putative regulatory protein
3.3	1.5	SCO6207	Uncharacterized protein
3.3		SCO5244	Anti-sigma factor antagonist PrsH
3.2		SCO1766	Putative glycohydrolase
3.2		SCO3484	Putative secreted sugar-binding protein
3.2		SCO2830	Probable amino acid ABC transporter protein
3.1		SCO5648	Fe(3+) ions import ATP-binding protein FbpC
3.0		SCO7292	Putative threonine dehydratase
3.0		SCO4585	Putative ABC transporter protein
3.0	2.8	SCO7325	Anti-sigma factor antagonist RsbV
2.9		SCO5289	Putative two component sensor kinase CvnA5
2.9		SCO1571	Uncharacterized protein
2.8		SCO3896	Putative RNA nucleotidyltransferase
2.8		SCO2123	Putative esterase/lipase
2.7		SCO0888	Putative secreted protein
2.7		SCO5701	Uncharacterized protein
2.7		SCO5240	Transcriptional regulator WhiB
2.6		SCO5367	ATP synthase AtpB
2.6		SCO5038	Putative integral membrane protein F42a
2.5		SCO4577	Putative helicase
0.4	0.5	SCO5290	Uncharacterized protein CvnB5
0.4		SCO0679	Uncharacterized protein
0.3		SCO3302	Putative integral membrane protein

Key:

5
4
3
2
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2

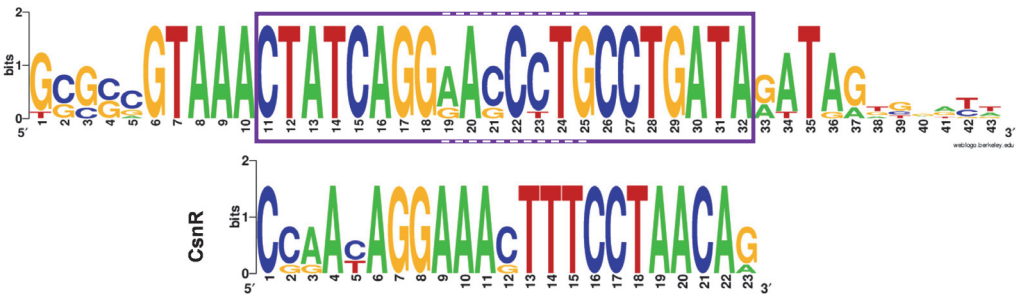
b)

Δ1447-wt		SCO#	Description
M	G		
10.4	6.0	SCO3967	Conserved hypothetical membrane protein
5.1	2.1	SCO1602	Uncharacterized protein
4.4	1.9	SCO5240	Transcriptional regulator WhiB
3.9	3.2	SCO2444	Putative fatty acid synthase
3.7	1.6	SCO2301	conserved hypothetical protein
3.7	2.0	SCO6068	Uncharacterized protein CvnB6
3.6	4.5	SCO0868	Putative regulatory protein
3.4		SCO5652	Uncharacterized protein
2.9		SCO5563	Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase ThiD
2.9	3.1	SCO4055	Putative alcohol dehydrogenase
2.8		SCO6207	Uncharacterized protein
2.6	1.9	SCO5024	Putative oxidoreductase
2.6	2.3	SCO6375	Putative secreted protein
2.5	1.5	SCO4444	Glutathione peroxidase
0.4	0.6	SCO2280	Putative transcriptional regulator
0.4	0.2	SCO5880	RedY protein
0.3	0.4	SCO5895	Putative methyltransferase RedI
0.2	0.2	SCO2256	3-methyl-2-oxobutanoate hydroxymethyltransferase PanB
0.1		SCO6804	Uncharacterized protein

FIGURE 3. Proteomic comparison of *S. coelicolor* M145 and its *rokL6* mutant.
Proteomic analysis was done for *S. coelicolor* M145 with or without added GlcN (a) and of *S. coelicolor* M145 and its *rokL6* deletion mutant (Δ1447) (b). Heat maps are shown of differentially expressed proteins and represent fold-changes, whereby changes of 2.5-fold or higher with a significance of $p < 0.1$ are shown. Fold changes are represented from red (0.2) to green (5.0) as shown in the colour key. See Table S3 in the supplemental material for the full data set. Strains were grown in liquid NMM media with mannitol and induced by adding of Mann or Mann + GlcN (50 mM each) for 1 h.

RokL6 and SCO1448, and eight orthologous pairs from other streptomycetes were analysed for conserved motifs using MEME. This identified a sequence of 38 nucleotides that is conserved in the intergenic region between *rokL6* and SCO1448 in *S. coelicolor* and other streptomycetes (Fig. 4A). The motif includes the inverted repeat CTATCAGG - 7 nt - CCTGATAG. The element was used to build a consensus sequence in PREDetector (Hiard *et al.*, 2007), which was then used to scan the *S. coelicolor* genome for similar motifs (Table S4). This identified three potential target sites with a medium score of 8 or higher in addition to the motif between *rokL6* and SCO1448. These putative binding sites were found in the regulatory region, upstream of gene SCO0317 encoding a putative transmembrane transport protein, SCO4114 encoding a sporulation-associated protein (Li *et al.*, 2006), and SCO2657, which encodes ROK-family transcriptional regulator CsnR. The latter is an important hit as CsnR regulates the uptake and metabolism of chitosan, the polymer of GlcN (Dubeau *et al.*, 2011; Viens *et al.*, 2015). None of the gene products of these genes were identified in the proteomic experiments so that at this stage no information on the effect on gene expression is available.

a)



b)

rokL6 -52bp- CTATCAGGCAGGCTCCCTGATAGTTTACGCCGCGGATTGACG -18bp-SCO1448

CTATCAGGCAGGCTCCCTGATAG	RokL6 (28)
CTATCAGGCAGGCTCCCTGATAG	CsnR (4.5)
GATAGTTTACGCCGCGGATTGAC	NagC (9.4)
ATAGTTTACGCCGCGGATTGACG	Mlc (5.4)

Figure 4. Predicted putative RokL6 binding site.
a) Schematic representation of MEME output for conserved DNA sequences identified upstream of *rokL6*. The consensus sequence was identified in the intergenic region of *rokL6*-SCO1448 and eight homologous pairs from other *Streptomyces* species using MEME. The 23 bp inverted repeat is highlighted. Also shown is the putative CsnR binding site (Dubeau et al., 2011).
b) Putative binding sites of ROK-family regulators RokL6, CsnR, NagC and Mlc in the intergenic region between *rokL6* and SCO1448 (top sequence), predicted by PREDetector, are shown. The regulators are listed on the right and the PREDetector score is given in brackets.

Interestingly, the predicted binding site of CsnR (Dubeau *et al.*, 2011) resembles that of RokL6 (Fig. 4A) and scanning the genome with this consensus sequence also identifies the same site upstream of *rokL6* (with a score of 4.5) as a potential CsnR target. The extensively studied ROK-family regulators Mlc and NagC from *E. coli*, involved in the regulation of aminosugar and glucose utilization, respectively, also have highly similar binding sites (Brechemier-Baey *et al.*, 2015). Indeed, scanning the genome of *S. coelicolor* using the consensus binding sequences of Mlc and NagC identified a putative site upstream of *rokL6* for both regulators, 1 nt apart (Fig. 4B). This putative binding site was located just upstream

of the predicted RokL6 site with a 5 nt overlap, suggesting some level of conservation in the recognition sites of ROK-family regulators. More extensive systems biology is required to definitively determine the binding site of RokL6, any cross-regulation with CsnR and establish the full extent of the RokL6 regulon.

THE *ROKL6-NAGB* DOUBLE MUTANT IS RESISTANT TO GLCN AND DOG.

As discussed in this Chapter and elsewhere in this thesis (Chapters II, IV, V and VII), a toxic molecule accumulates in *nagB* null mutants when grown in the presence of either GlcN or GlcNAc. However, the nature of this molecule is as of yet unknown. A well-known toxic compound is 2-deoxyglucose (DOG), a non-metabolisable glucose derivative with a hydrogen in the place of its 2-hydroxyl group, that inhibits growth of prokaryotic and eukaryotic cells. In addition to applications as a glycolytic inhibitor, DOG is also being investigated as an anticancer drug (Zhang *et al.*, 2014; Bost *et al.*, 2016). In most organisms, DOG is taken up via the native glucose import system and subsequently phosphorylated to 2-deoxyglucose 6-phosphate (DOG-6P). Challenge of streptomycetes with the toxic molecule results in spontaneous *glk* mutants, lacking glucose kinase activity (Ikeda *et al.*, 1984). Though DOG-6P is responsible for the inhibition of glycolysis, by interfering with phosphoglucose isomerase (PGI) activity (Wick *et al.*, 1957), this is insufficient to cause the growth inhibition observed in the presence of DOG (Ralser *et al.*, 2008).

Since glucose and aminosugar metabolism are closely linked, we tested whether any of the mutants created in this study might override toxicity to DOG by testing for their ability to grow on MM with 50 mM DOG (Fig. 5). As expected, growth of wild-type *S. coelicolor* was inhibited by DOG, as were mutants deleted for *nagB*, *rokL6,* SCO1448, SCO6110, *nagA* and SCO4393 (the latter two not shown). Intriguingly, the *rokL6-nagB* double mutant showed significant DOG resistance. This suggests that RokL6 may play an important role in the metabolism of DOG, and in particular in the accumulation of a toxic compound derived thereof.

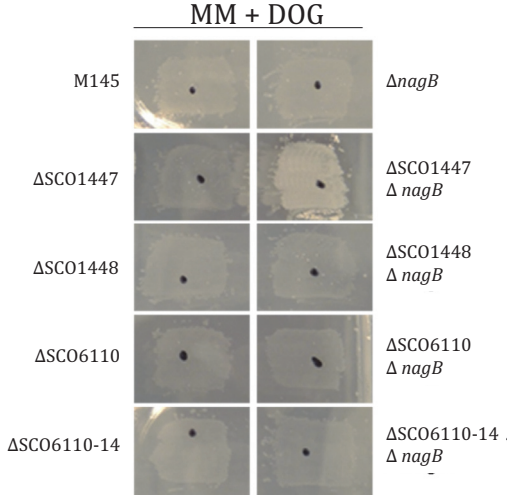


FIGURE 5. Phenotypic analysis of *S. coelicolor* mutants on 2-deoxy-D-glucose (DOG).
Phenotypic analysis of *S. coelicolor* M145 and deletion mutants (Δ) as annotated. Spores were plated onto minimal medium (MM) supplemented with 50 mM DOG.

Discussion

The aminosugar *N*-acetylglucosamine (GlcNAc) is a preferred nutrient for *Streptomyces* and also acts as a signalling molecule for the nutritional status of the environment. Antibiotic production and development are activated by GlcNAc under poor growth conditions (famine), while these processes are blocked under rich conditions (feast) (Rigali *et al.*, 2008). This property of GlcNAc has been exploited to screen for novel antibiotics (Zhu *et al.*, 2014a). While GlcNAc metabolism and transport in streptomycetes are generally well understood, little was known about how the de-acetylated form, GlcN, is metabolised in these organisms. Our initial experiments suggested that GlcN may be metabolised via the GlcNAc pathway, and among others revealed the likely involvement of the ROK-family protein RokL6 (Świątek, 2012; Świątek *et al.*, 2012b; Chapter V).

The impact of GlcN induction was examined at the level of protein expression. GlcN induced the expression of PTS GlcNAc-specific membrane component NagE2 over five-fold in *S. coelicolor* (Nothaft *et al.*, 2010). GlcN-6P is the ligand for DasR, and renders its DNA-binding activity inactive (Rigali *et al.*, 2006). Therefore, the effect of GlcN on NagE2 expression may be governed indirectly, via the inhibition of DasR, which represses the transcription of *nagE2*. However, we cannot rule out that NagE2 is involved in GlcN transport, especially since many of the known DasR targets did not show elevated protein levels under the same growth conditions. The expression of SCO3484, SCO2830 and SCO4585 was also strongly increased when GlcN was added, and both are ATP-binding proteins of ABC transporters. SCO3484 is strongly repressed by DasR (Świątek-Połatyńska *et al.*, 2015), and may therefore be induced via GlcN-mediated metabolic inactivation of DasR. SCO4585 is part of an ABC-type multidrug transport system and SCO2830 is predicted to be involved in amino acid transport.

To obtain further insights into the role of RokL6 in the control of gene expression in *S. coelicolor*, the *rokL6* deletion mutant was also examined by proteome analysis. The deletion of *rokL6* did not significantly influence genes involved in aminosugar metabolism and GlcN sensing. The absence of RokL6 resulted in a significant decrease in levels of proteins related to prodigiosin production, which suggests that RokL6 is directly or indirectly involved in the control of prodiginine biosynthesis. Whether RokL6 directly regulates secondary metabolism remains to be determined through DNA binding studies. The data suggest that RokL6 could also have a role in development; the levels of sporulation regulator WhiB were strongly increased in the absence of RokL6. Additionally, a putative RokL6 binding site was predicted upstream of the gene encoding sporulation-associated protein SCO4114 (Li *et al.*, 2006).

A 23bp inverted repeat (CTATCAGG - 7 nt – CCTGATAG) was identified in the intergenic region between *rokL6* and SCO1448. Regulatory proteins often autoregulate their own gene expression, therefore this element qualifies as a potential RokL6 binding site. The element was used to scan the *S. coelicolor* genome for similar motifs using the PREDetector algorithm (Hiard *et al.*, 2007). In addition to the putative auto-regulatory site, a motif was also identified upstream of SCO0317 encoding a transporter with unknown substrate. Interestingly, a putative binding site was also identified upstream of *csnR*, which controls utilization of the GlcN polymer chitosan scanning the genome with the CsnR element identified the same 23bp inverted repeat as a potential binding site. The extensively studied ROK-family regulators Mlc and NagC from *E. coli*, involved in the regulation of aminosugar and glucose utilization, respectively, also have highly similar binding sites. However, it was shown that under native conditions the regulators bind exclusively to their own targets, demonstrating the subtlety and fine-tuning in the regulatory function of these proteins

(Brechemier-Baey *et al.*, 2015). Systems biology should reveal the binding site for RokL6 and establish the full extent of the RokL6 regulon.

A major question to answer is, what is the toxic molecule that accumulates in *nagB* mutants grown on aminosugars? An important cue may come from the fact that *rokL6-nagB* double mutants can grow on 2-deoxyglucose (DOG), while single mutants deleted for either *rokL6* or *nagB* cannot. This connects glucose and aminosugar metabolism. DOG is a toxic analogue of glucose, which is also a safe and effective chemotherapeutic agent against various cancers (Thompson & Kleinzeller, 1989; Zhang *et al.*, 2014). DOG is a known inhibitor of glycolysis; however, this does not fully explain the growth inhibition in eukaryotic cells (yeast and mammalian cells) (Ralser *et al.*, 2008; Thompson & Kleinzeller, 1989). Our data suggest that the toxicity may instead be mediated by an intermediate or end product of aminosugar metabolism and that this may directly connect to the toxicity of DOG. The surprising ability of the *rokL6-nagB* mutant to grow on DOG suggests that DOG toxicity is mediated via aminosugar metabolism. This discovery is not only relevant for the understanding of the function of RokL6, but may also provide new insights into the mode of action of the important anticancer drug 2-deoxyglucose.

Taken together, our studies on transcriptional repressor RokL6 shed new light on aminosugar metabolism in *Streptomyces*. Like for most ROK-family regulators, the ligand and regulon for RokL6 remain to be identified. Our data will help define a clearer and more concise picture of the pathways for aminosugar metabolism, specifically GlcN, as well as the intricate deviations and overlaps between GlcNAc, GlcN and potentially DOG metabolism and regulation.

ACKNOWLEDGEMENTS

We are very grateful to Chao Du for proteomics support.