Exploring the metabolism and toxicity of amino sugars and 2-deoxyglucose in Streptomyces

Li, C.

Chapter 3

Systems-wide analysis of the ROK-family regulatory gene *rokL6* and its role in the control of glucosamine toxicity in *Streptomyces coelicolor*

Chao Li[a], Mia Urem[b], Chao Du[a], Le Zhang[a] and Gilles P. van Wezel[a,*]

Published as:
Li, C., Urem, M., Du, C., Zhang, L., and van Wezel, G.P. (2023) Systems-wide analysis of the ROK-family regulatory gene *rokL6* and its role in the control of glucosamine toxicity in *Streptomyces coelicolor. Appl Environ Microbiol* **89**: e01674-01623.

Abstract

Streptomycetes are saprophytic bacteria that grow on complex polysaccharides, such as cellulose, starch, chitin, and chitosan. For the monomeric building blocks glucose, maltose, and *N*-acetylglucosamine, the metabolic pathways are well-documented, but that of glucosamine (GlcN) is largely unknown. *Streptomyces nagB* mutants, which lack glucosamine-6-phosphate deaminase activity, fail to grow in the presence of high concentrations of GlcN. Here, we report that mutations in the gene for the ROK-family transcriptional regulator RokL6 relieve the toxicity of GlcN in *nagB* mutants, as a result of elevated expression of the major facilitator superfamily (MFS) exporter SCO1448. Systems-wide analysis using RNA sequencing, ChIP-Seq, EMSAs, 5′RACE, bioinformatics, and genetics revealed that RokL6 is an autoregulator that represses the transcription of SCO1448 by binding to overlapping promoters in the *rokL6*-SCO1448 intergenic region. RokL6-independent expression of SCO1448 fully relieved the toxicity of GlcN to *nagB* mutants. Taken together, our data show a novel system of RokL6 as a regulator that controls the expression of the MFS transporter SCO1448, which in turn protects cells against GlcN toxicity, most likely by exporting toxic metabolites out of the cell.

Introduction

Streptomycetes are Gram-positive bacteria with a mycelial lifestyle that reproduce via sporulation. Their large GC-rich genomes encode a plethora of specialized metabolites such as antibiotics, anticancer drugs, and many other industrially and medically relevant compounds (Berdy, 2005, Hopwood, 2007, Barka *et al.*, 2016). The production of these molecules is closely related to the transition from vegetative to aerial growth during the development of the colonies (Bibb, 2005, Van Der Heul *et al.*, 2018). Streptomycetes grow by tip extension and branching of vegetative hyphae, which are divided into multigenomic compartments. Under adverse conditions such as nutrient depletion, a complex developmental program is initiated, which results in the formation of an aerial mycelium, and eventually the aerial hyphae differentiate to form chains of unigenomic spores (Barka *et al.*, 2016, Flärdh & Buttner, 2009). At the onset of development, the vegetative mycelium is partially degraded via programmed cell death so as to provide the building blocks necessary for aerial growth in an otherwise nutrient-depleted environment (Manteca *et al.*, 2005b, Tenconi *et al.*, 2018).

During cell wall recycling, the constituents 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. The lactyl ether substituent of the intracellular MurNAc-6-phosphate (MurNAc-6P) is cleaved by MurNAc-6P etherase, yielding *N*-acetylglucosamine-6-phosphate (GlcNAc-6P) (Borisova *et al.*, 2016). GlcNAc is thereby internalized by the phosphoenolpyruvate-dependent phosphotransferase system (PTS), which simultaneously phosphorylates GlcNAc to GlcNAc-6P (Nothaft *et al.*, 2003a, Nothaft *et al.*, 2010). The next step is GlcNAc-6P deacetylation by *N*-acetylglucosamine-6-phosphate deacetylase NagA forming glucosamine-6-phosphate (GlcN-6P), which is a central molecule at the intersection of multiple metabolic pathways, including glycolysis via conversion to fructose-6-phosphate (Fru-6P) by glucosamine-6-phosphate deaminase NagB (Świątek *et al.*, 2012a). GlcNAc plays a critical role in signalling during the control of development and antibiotic production in streptomycetes (Rigali *et al.*, 2018, Urem *et al.*, 2016). DasR controls amino sugar transport and metabolism and is also a highly pleiotropic repressor of antibiotic production in *Streptomyces* (Rigali *et al.*, 2006, Rigali *et al.*, 2008, Świątek-Połatyńska *et al.*, 2015). Metabolic control of DasR is a key step in the early activation of development and specialized metabolism under nutrient-deprived conditions.

While the metabolic pathway of GlcNAc has been well-characterized, little is known of glucosamine (GlcN) metabolism in streptomycetes. The hydrolysis of chitosan and the metabolism of chitosan-derived oligomers $(GlcN)_{2\text{-}3}$ are under the control of CsnR (Dubeau *et al.*, 2011, Viens *et al.*, 2015), a repressor from the ROK [repressors, open reading frames (ORFs), and kinases] family of transcriptional regulators (Titgemeyer *et al.*, 1994). GlcN oligomers are imported via the ABC-transporter complex CsnEFG-MsiK and are then

presumably hydrolyzed and phosphorylated by CsnH, a sugar hydrolase, and CsnK, a ROK-family kinase, respectively (Viens *et al.*, 2015).

High concentrations of either GlcN or GlcNAc are toxic to *Streptomyces coelicolor* in the absence of the enzyme glucosamine-6-phosphate deaminase (NagB) (Świątek *et al.*, 2012a). Similar amino sugar sensitivity was also observed for *Escherichia coli* (Plumbridge, 2009, Plumbridge, 2015). Despite extensive studies focusing on the metabolism of amino sugars, the underlying cause of amino sugar toxicity remains unresolved. Interestingly, when grown in the presence of high concentrations of either GlcN or GlcNAc, *S. coelicolor* Δ*nagB* strains sustain spontaneous second-site mutations that allow the colonies to survive. We previously exploited this principle to identify novel genes related to GlcN transport and metabolism (Świątek *et al.*, 2012a, Świątek *et al.*, 2012b). Spontaneous mutations or deletion of *nagA*, preventing the conversion from GlcNAc-6P to GlcN-6P, relieves the toxicity of both GlcNAc and GlcN to *S. coelicolor nagB* mutants (Świątek *et al.*, 2012b). Since NagA is not known to play a role in GlcN metabolism, this points to a caveat in our understanding of amino sugar metabolism. In addition to suppressor mutations in *nagA*, two independent suppressor mutations were found in the gene for ROK-family regulator SCO1447 (RokL6), which relieved toxicity specifically of GlcN but not of GlcNAc. ROK-family of proteins often play important roles in the (control of) sugar utilization in bacteria, and these regulators are widespread in streptomycetes (Titgemeyer *et al.*, 1994, Świątek *et al.*, 2013, Lu *et al.*, 2020).

Here, we report on the function of RokL6 by a systems-wide approach, using mutational and transcriptional analysis and *in vitro* and *in vivo* DNA binding studies. This revealed that the gene product of *rokL6* is an ROK-family regulator that specifically represses the adjacent gene SCO1448. This gene encodes a putative major facilitator superfamily (MFS) transporter, and we propose that it acts as a pump for toxic substances created during the challenge of *nagB* mutants with higher concentrations of GlcN.

Materials and Methods

Strains, plasmid, and growth conditions

The bacterial strains and plasmids used or constructed in this study are summarized and listed in **Table S1**, and all oligonucleotides are described in **Table S2**. *Escherichia coli* was grown and transformed according to the standard procedures (Sambrook *et al.*, 1989), with *E. coli* JM109 serving as the host for routine cloning, and *E. coli* ET12567 (MacNeil *et al.*, 1992) used for the isolation of non-methylated DNA for transformation into *Streptomyces*. For protein heterologous expression, *E. coli* BL21(DE3) from Novagen was used. *E. coli* was grown in Luria-Bertani media in the presence of selective antibiotics as required, with the following final concentrations: ampicillin (100 μg·ml^{-1}), apramycin (50 μg·ml^{-1}), kanamycin (50 μg·ml^{-1}), and chloramphenicol (25 μg·ml^{-1}). *Streptomyces coelicolor* M145 (Hoskisson & van Wezel, 2019)

was obtained from the John Innes Centre strain collection and was the parent of all mutants. *S. coelicolor nagB* mutant Δ*nagB* (Świątek *et al.*, 2012b) and GlcN-derived *nagB* suppressor mutant SMG1 (Swiatek, 2012) have been described previously. All *Streptomyces* media and routine techniques are based on the *Streptomyces* manual (Kieser *et al.*, 2000). Phenotypic characterization of *Streptomyces* mutants was carried out on minimal medium (MM) agar plates with different carbon sources as indicated. Soy flour mannitol (SFM) agar plates were used for the preparation of spore suspensions and *E. coli* to *S. coelicolor* conjugations. A mixture of 1:1 yeast-extract malt extract and tryptic soy broth liquid media was used to cultivate mycelia for protoplast preparation and genome DNA isolation. Growth curves of *Streptomyces* were performed in liquid cultures containing minimal medium normal minimal media phosphate (NMMP) (Kieser *et al.*, 2000) with 1% (wt/vol) glucose or GlcN as the sole carbon source, and the dry weights of the mycelia were measured at different time points.

Gene knock-out, complementation, and overexpression

As a basis for the creation of gene deletion mutants, we used the unstable multi-copy plasmid pWHM3 (Vara *et al.*, 1989, Zhang *et al.*, 2018), as described previously (Świątek *et al.*, 2012a). For the *rokL6* knock-out construct, a 1,254-bp 5′ flanking region and a 1,366-bp 3′ flanking region were amplified by PCR from the *S. coelicolor* M145 genome, using the primer pairs described in **Table S2**. The upstream region was cloned as an *Eco*RI-*Xba*I fragment, and the downstream region as an *Xba*I-*Bam*HI fragment, and these two fragments were ligated into pWHM3 from the *Eco*RI and *Bam*HI sites. The apramycin resistance cassette *aac(3)IV*, flanked by *lox*P sites (apra-*lox*P), was subsequently cloned into the engineered *Xba*I sites between the flanks to create *rokL6* knock-out construct pKO-*rokL6*. In the same way, the flanking regions of SCO1448 were cloned into pWHM3 with the apra-*lox*P to produce the SCO1448 knock-out vector, pKO-1448. The knock-out constructs were introduced into *S. coelicolor* M145 or its *nagB* mutant. For clean gene knock-out mutants, the apramycin resistance cassette was excised by the introduction of pUWLcre, which expresses Cre recombinase (Fedoryshyn *et al.*, 2008, Khodakaramian *et al.*, 2006). The correct recombination event in each of the knock-out mutants was confirmed by PCR.

For *rokL6* complementation, the entire coding region of *rokL6* with its own promoter was amplified from the *S. coelicolor* chromosome. The PCR product was digested with *Xba*I/*Eco*RV and then inserted into pSET152 (Bierman *et al.*, 1992) to obtain *rokL6*-complemented vector pCOM-*rokL6*. This construct was introduced into *nagB*-*rokL6* double mutant Δ*nagB*Δ*rokL6*, and the successfully complemented strains Δ*nagB*Δ*rokL6*-C were selected by apramycin. The empty pSET152 was conjugated in Δ*nagB*Δ*rokL6* to produce Δ*nagB*Δ*rokL6*-E, which was used as a control strain.

For SCO1448 overexpression, a 1,212-bp DNA fragment containing SCO1448 was amplified and ligated into pSET152 with the 63-bp highly efficient promoter P30, namely the

promoter SP30 with the 20-bp ribosomal binding site RBS15 (Bai *et al.*, 2015), to produce SCO1448-overexpressing vector pOE-1448. The pOE-1448 construct was then introduced into Δ*nagB* to obtain the SCO1448 overexpression strain Δ*nagB*SCO1448-O.

$FLAG_3$ tag knock-in by CRISPRTo express FLAG-tagged RokL6 in *S. coelicolor*, the 3× FLAG sequence ($FLAG_3$) was fused to the end of the original copy of *rokL6* on the genome using codon-optimized CRISPR-Cas9 system (Cobb *et al.*, 2015) as described previously (Du *et al.*, 2022). Briefly, the spacer sequence (5′- GGCGCCGCGGGCTACCGCTC-3′) specific to *rokL6*, located at the end of the coding region, was inserted into the pCRISPomyces-2 plasmid from *Bbs*I sites. Next, the 2,263-bp template containing $FLAG_3$ for homology-directed repair was made and inserted into the spacer-containing plasmid, resulting in a *rokL6*- $FLAG_3$ knock-in construct designated as pKI-$FLAG_3$. Mutagenesis was done according to the previous study (Du *et al.*, 2022). After conjugation of pKI-$FLAG_3$ to *S. coelicolor* A3(2) M145, ex-conjugants were patched on SFM agar plates with 20 $\mu g \cdot ml^{-1}$ apramycin. Then, positive ex-conjugants were patched on antibiotic-free SFM agar plates and grown at 37°C. Spores were collected and checked for loss of construct. Apramycin-sensitive strains were selected for spore collection, and their genomes were checked for desired recombination events. The successful *in situ* $FLAG_3$ knock-in strain was identified as M145-*rokL6*-FLAG. Additionally, *rokL6* with $FLAG_3$ was ligated into pSET152 to generate pCOM-$FLAG_3$ and introduced into the *nagB*-*rokL6* double mutant for evaluating RokL6-$FLAG_3$ function *in vivo*.

Heterologous expression and purification of His_6-tagged RokL6 protein

For the heterologous expression of *S. coelicolor* RokL6 in *E. coli*, the 1,197-bp *rokL6* coding region was amplified from *S. coelicolor* genomic DNA using primer pair RokL6-exp-F/RokL6-exp-R. The PCR fragment was ligated into pET-28a (+) from *Xho*I and *Nco*I sites, generating expression vector pEX-RokL6. The RoKL6 expression vector was transformed into *E. coli* BL21(DE3), and the expression of C-terminal His_6-tagged RokL6 recombinant protein, RokL6-His_6, was induced by the addition of isopropyl β-D-1-thiogalactopyranoside at a final concentration of 1 mM when the cell density was reached around an optical density at 600 nm of 0.6, followed by overnight incubation at 16°C. Cells were harvested, washed, and disrupted in lysis buffer (Guo *et al.*, 2018) by sonication. Soluble RokL6-His_6 was purified from the supernatant using HisPur Cobalt Resin (Thermo Fisher Scientific, USA) and dialyzed against electrophoretic mobility shift assay (EMSA) binding buffer.

Chromatin immunoprecipitation sequencing of RokL6

Chromatin immunoprecipitation sequencing (ChIP-Seq) experiments were carried out essentially as described previously (Du *et al.*, 2022). In brief, *S. coelicolor* M145-*rokL6*-FLAG was grown on MM agar covered with cellophane disks, using mannitol with and without 50 mM GlcN as the carbon source. Mycelia were collected after 24 h (vegetative growth) and 48

h (sporulation). Mycelia were treated with phosphate-buffered saline (PBS) buffer containing 1% formaldehyde for 20 min to cross-link the DNA and protein. After thorough washing in PBS, mycelia were resuspended in lysis buffer [10 mM Tris-HCl pH 8.0, 50 mM NaCl, 15 $mg \cdot ml^{-1}$ lysozyme, 1× protease inhibitor (Roche, Bavaria, Germany)] and incubated at 37°C for 20 min. After incubation, 0.5 ml IP buffer (100 mM Tris-HCl pH 8.0, 250 mM NaCl, 0.8%, vol/vol Triton-X-100) was added to the mycelia samples, and chromosomal DNA was sheared to 100–500 bp fragments using the Bioruptor Pluswater bath sonication system (Diagenode, Liège, Belgium). The lysates were incubated with 40 µl Anti-FLAG M2 affinity gel (cat A2220, Sigma-Aldrich, St. Louis, USA) and incubated at 4°C overnight. After centrifugation, the pellet and untreated total extracts (control) were incubated in 100 µl of IP elution buffer (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 1%, [wt/vol] SDS) at 65°C overnight to reverse the cross-link. The beads were removed by centrifugation before DNA isolation by phenol-chloroform. The extracted DNA samples were then further purified with the DNA Clean & Concentrator kit (Zymo Research, CA, USA). The enriched DNA samples were sent to Novogene Europe (Cambridge, UK) for library construction and next-generation sequencing. ChIP-Seq data analysis was performed as described previously (Du *et al.*, 2022).

RNA isolation, RNA sequencing, and quantitative PCR

Spores (10^7 CFU) of *S. coelicolor* M145 and the *rokL6* deletion mutant were grown on MM agar plates overlayed with cellophane discs, with either 1% (wt/vol) mannitol or 1% mannitol + 50 mM GlcN as the carbon sources. Biomass from two time points (24 h on mannitol or 26 h on mannitol with GlcN for vegetative growth phase, VEG; and 42 h on mannitol or 44 h on mannitol with GlcN for sporulation phase, SPO) was collected and snap-frozen in liquid N_2. All samples were assessed as biological triplicates. After breaking the mycelia using a TissueLyser II (Qiagen, Venlo, The Netherlands), the RNA was extracted using a modified Kirby mix (Kieser *et al.*, 2000). The transcriptome sequencing library preparation and sequencing were outsourced to Novogen Europe (Cambridge, UK). Removal of rRNA from the samples was carried out using NEBNext Ultra directional RNA Library Prep Kit (NEB, MA, USA). Sequencing libraries were generated using the NEBNext Ultra RNA Library Prep Kit for Illumina (NEB), and sequencing was performed on an Illumina NovaSeq 6000 platform. Raw data were cleaned using fastp v0.12.2 (Chen *et al.*, 2018), and then mapped to the *S. coelicolor* M145 genome sequence (GenBank accession AL645882.2) using bowtie2 v2.4.4 (Langmead & Salzberg, 2012). Read counts for each gene were generated by featureCounts v2.0.1 (Liao *et al.*, 2014). Values for transcripts per million were generated using a custom Python script. Differentially expressed genes and log2Foldchange were determined using DESeq2 v1.32.0 (Love *et al.*, 2014) with the data shrinkage function “apeglm” (Zhu *et al.*, 2019).

For quantitative PCR (qPCR), cDNA was synthesized using the iScript cDNA Synthesis Kit (Biorad, CA, USA). Briefly, qPCR was performed using the iTaq Universal SYBR green qPCR Kit

(Biorad, CA, USA), and the program used was set as follows: 95°C for 30 s; 40 cycles of 95°C for 10 s, 60°C for 30 s, plate reading; melting curve from 65°C to 95°C with 5 s per 0.5°C increment. The principal RNA polymerase σ factor encoding gene, *hrdB* (SCO5820), was used as the internal control, and the qPCR data were analysed by CFX Manager software (version 3.1, Biorad, CA, USA), using ΔΔCq standard, which is an implementation of the method described in reference (Vandesompele *et al.*, 2002).

Electrophoretic mobility shift assays

Electrophoretic mobility shift assays (EMSAs) were performed as described previously (Du *et al.*, 2022). Double-stranded DNA probes (60 bp) were generated by gradually cooling reverse complemented single-strand oligonucleotides in 30 mM HEPES, pH 7.8, heating to 95°C for 5 min, then ramping to 4°C at a rate of 0.1°C·s^{-1}. The *in vitro* DNA-protein binding assays were performed in EMSA binding buffer (20 mM HEPES pH 7.6, 30 mM KCl, 10 mM $(NH_4)_2SO_4$, 1 mM EDTA, 1 mM DTT, 0.2% Tween20). The binding reactions (10 µl), including 1 picomole DNA probe, 20 ng·$µl^{-1}$ bovine serum albumin (BSA), and various concentrations of RokL6-His_6, were incubated at 30°C for 20 min. The reactions were then loaded on 5% non-denatured polyacrylamide gels and separated by electrophoresis. The gels were briefly stained with ethidium bromide and imaged using the Gel Doc imaging system (BioRad, CA, USA).

Determination of transcriptional start sites

Transcriptional start sites (TSS) of *rokL6* and SCO1448 were determined by 5′ RACE using a 5′ /3′ RACE Kit 2nd generation (Roche, CA, USA). Total RNA (2 µg) extracted from 48-h cultures of *S. coelicolor* grown on solid minimal medium containing 1% mannitol was used for reverse transcription with gene-specific primers. The obtained cDNA was purified, and an oligo(dA) tail was added to the 3′ end by terminal transferase, followed by PCR amplification of the tailed cDNA with oligo(dT) anchor primer and gene-specific nested primers. Using the resulting PCR product (diluted 1,000-fold) as a template, an additional round of PCR was performed with a more inner nested primer and an anchor primer provided in the kit to produce a single specific DNA band. The final PCR product was purified and sent for sequencing, and the TSS was determined as the first nucleotide following oligo(dA).

Promoter activity test

The promoter regions of *rokL6*, SCO1448, *nagKA* (Świątek *et al.*, 2012a), and P30 (Bai *et al.*, 2015) were ligated with a gene expressing enhanced green fluorescent protein (eGFP) and integrated into pSET152 via the *Xba*I/*Eco*RV sites, generating four recombinant vectors: pEGFP-*rokL6*, pEGFP-1448, pEGFP-*nagKA*, and pEGFP-P30. The recombinant constructs were then introduced into *S. coelicolor* and *rokL6* deletion mutant Δ*rokL6* to generate the following

strains: M145-P*rokL6*-eGFP, M145-P1448-eGFP, M145-P*nagKA*-eGFP, M145-P30-eGFP, and Δ*rokL6*-P*rokL6*-eGFP, Δ*rokL6*-P1448-eGFP, Δ*rokL6*-P*nagKA*-eGFP, and Δ*rokL6*-P30-eGFP. The *in vivo* activities of all the promoters were evaluated by comparing the transcription level of the gene for eGFP via confocal microscopy and confirmed by qPCR.

Confocal imaging

Sterile coverslips were inserted into MM with 1% mannitol agar plates at an angle of 45°, and spores of eGFP-harbouring strains were inoculated at the intersection angle and incubated at 30°C for 48 h. The mycelium was scraped off from coverslips into a drop of water and then imaged with a TCS SP8 confocal microscope (Leica) using a 63× oil immersion objective (NA: 1.40) (Celler *et al.*, 2016). Image processing and the analysis of fluorescent intensity were performed using ImageJ (version 1.54d).

Bioinformatics analysis

DNA and protein database searches were performed using the BLAST server of the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov) and the *S. coelicolor* genome page services (http://strepdb.streptomyces.org.uk). Polypeptide sequences were aligned using Clustalw available at http://www.ebi.ac.uk/clustalw. Gene synteny analysis was performed using SynTax (Oberto, 2013). To visualize the consensus sequence for the predicted binding site of RokL6, WebLogo (Crooks *et al.*, 2004) was used, and regulon predictions were done using PREDetector (Hiard *et al.*, 2007). The comparative analysis of the intergenic regions of *rokL6*/SCO1448 and their orthologous pairs was performed by MEME (Bailey *et al.*, 2009).

Data availability

Clean RNA-Seq reads and gene read-counts tables are available at GEO database (Barrett *et al.*, 2012) with accession number GSE234437. Clean ChIP-Seq reads and binding region identification (peak calling) files are available at GEO database with accession number GSE234438.

Results

RokL6 is involved in GlcN metabolism

We previously showed that null mutants of *S. coelicolor nagB* are sensitive to high concentrations of GlcN, whereby spontaneous second-site mutations arise that allow the colonies to survive. Exploitation of this principle led to the identification of suppressor mutations in *nagA* and in a novel gene related to GlcN transport and metabolism (Świątek *et al.*, 2012a, Świątek *et al.*, 2012b). Two independent suppressor mutants were obtained in the gene SCO1447 when *S. coelicolor* Δ*nagB* was grown on GlcN, namely SMG1 and SMG38, and

these mutations specifically alleviated the toxicity exclusively on GlcN, but not GlcNAc (Świątek *et al.*, 2012b, Urem, 2017). Suppressor mutant SMG1 had sustained a single nucleotide insertion at nucleotide (nt) position 26 within the coding region of SCO1447, while SMG38 had a single nucleotide deletion at nt position 120 in SCO1447. In both cases, the mutations resulted in a frameshift at the beginning of the gene, thereby preventing the expression of the active protein. SCO1447 encodes a ROK-family transcriptional regulator, which we designated RokL6, based on the terminology that ROK-family proteins are named after the specific cosmid they are located on (Świątek *et al.*, 2013) within the ordered cosmid library that was used for the *S. coelicolor* genome sequencing project (Redenbach *et al.*, 1996).

The observation that the mutations specifically alleviate the toxicity of GlcN to *nagB* null mutants, but not that of GlcNAc, suggests that RokL6 plays a specific role in the control of GlcN metabolism. To test this, the *rokL6* single mutant Δ*rokL6* and *nagB-rokL6* double mutant Δ*nagB*Δ*rokL6* were generated by deleting *rokL6* in *S. coelicolor* M145 and in the previously published *nagB* null mutant (Swiatek, 2012), respectively. To create the mutants, the entire coding region of *rokL6* (nt positions + 3 to +1,200) was replaced with the apramycin-resistance cassette *aac(3)IV*, which was flanked by *lox*P sites. The *aac(3)IV* gene was subsequently excised from the genome using the Cre recombinase expressed from plasmid pUWLcre. The resulting *rokL6* single mutant and *nagB-rokL6* double mutant, Δ*nagB*Δ*rokL6*, were confirmed by PCR. In order to ascertain that the phenotypes were specifically caused by the deletion of *rokL6*, we genetically complemented the Δ*nagB*Δ*rokL6* mutant by expressing *rokL6*. For this, the −478/+1,226 region of *rokL6*, harboring the entire gene and its promoter region, was amplified from the *S. coelicolor* chromosome and cloned into integrative vector pSET152 to obtain pCOM-*rokL6* (see **Materials and Methods** for details).

As expected, *S. coelicolor* M145 and its *rokL6* single mutant Δ*rokL6*, which has an intact copy of *nagB*, grew well on all media, while *nagB* mutants were unable to grow on MM with mannitol (1%, wt/vol) and 5 mM GlcN or GlcNAc. The *rokL6-nagB* double mutant grew well on MM with 1% mannitol and 5 mM GlcN, but failed to grow when GlcNAc was used instead of GlcN (**Fig. 1**). Importantly, transformants expressing RokL6 in the Δ*nagB*Δ*rokL6* double mutant via the introduction of pCOM-*rokL6* failed to grow, similar to the *nagB* mutant. These data strongly suggest that the deletion of *rokL6* was the sole reason why Δ*nagB*Δ*rokL6* double mutants could grow on GlcN (**Fig. 1**).

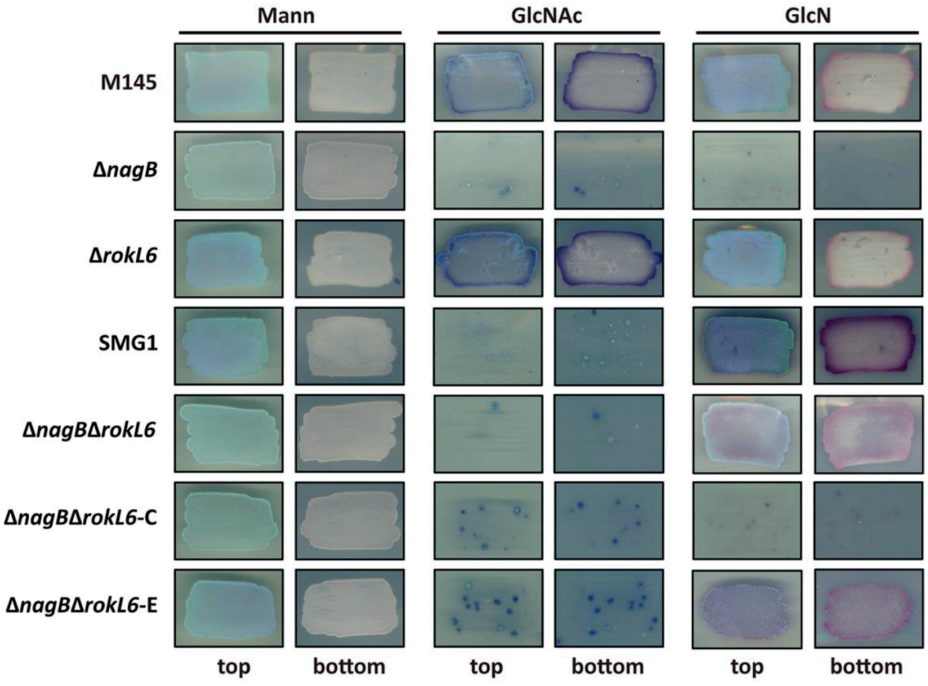

Figure 1. Sensitivity of *S. coelicolor* mutants to GlcN and GlcNAc. Spores (10^5 CFU) of *S. coelicolor* M145 or its mutant derivatives were streaked onto minimal medium (MM) with either 1% (w/v) mannitol (Mann), 5 mM *N*-acetylglucosamine (GlcNAc) or 5 mM glucosamine (GlcN) and grown for 72 h at 30°C. Strains were *S. coelicolor* M145 (M145), its mutant derivatives Δ*nagB*, Δ*rokL6*, suppressor mutant SMG1, Δ*nagB*Δ*rokL6*, and the Δ*nagB*Δ*rokL6* mutant harboring either pCOM-*rokL6* (Δ*nagB*Δ*rokL6*-C) or empty vector pSET152 (Δ*nagB*Δ*rokL6*-E). Both top and bottom view are shown. Note that all strains without the gene *nagB* are sensitive to GlcNAc, while all strains lacking *rokL6* are resistant to GlcN.

Conservation and gene synteny of *rokL6*

The gene *rokL6* (SCO1447) from *S. coelicolor* consists of 1,200 nucleotides and encodes a protein of 399 amino acids. RokL6 is characterized by an N-terminal winged helix-turn-helix DNA-binding site, which is found in various families of DNA-binding proteins and a sugar kinase domain annotated as a putative ROK-family regulator. Protein alignment shows a high amino acid sequence conservation of RokL6 in *Streptomyces* species (**Fig. S1**). Its genomic neighbors, SCO1446 and SCO1448, encode a putative integral membrane protein and an MFS transporter with unknown substrates, and the intergenic regions of *rokL6*-SCO1446 and *rokL6*-SCO1448 are 21- and 112-bp long, respectively (**Fig. 2A**). There is significant gene synteny for the genomic region surrounding *rokL6* and its orthologs in *Streptomyces*. In particular, SCO1448 and its homologs often lie divergently transcribed from the corresponding ROK regulator

encoding genes (**Fig. 2B**). The *rokL6*-SCO1448 unit is also identified in many *Kitasatospora* species (**Fig. S2**).

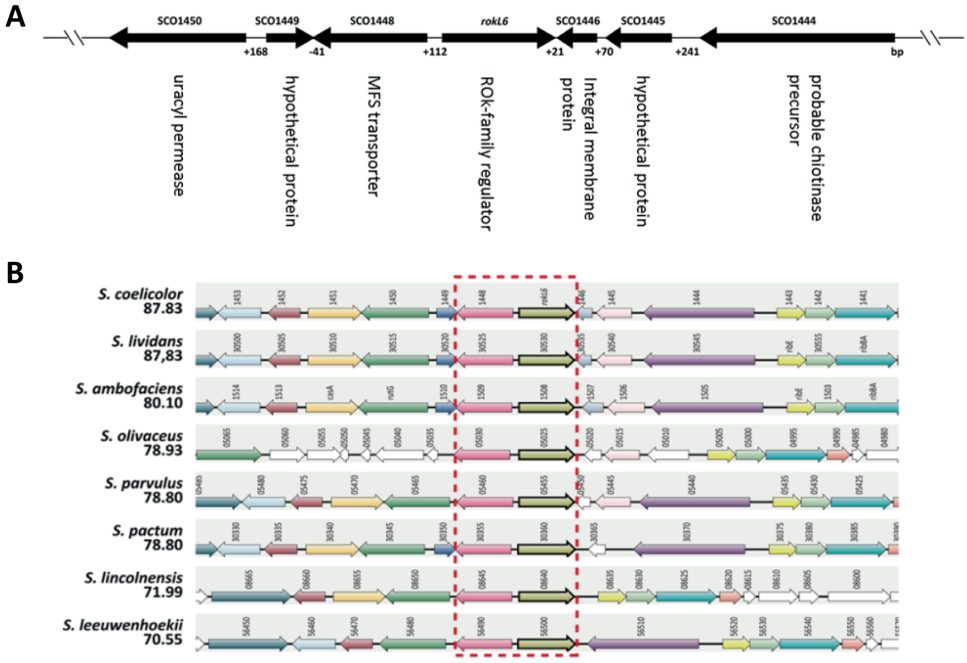

Figure 2. Genetic organization and gene synteny around *rokL6*. (**A**) Genetic organization of the genomic region around *rokL*6. All open reading frames (ORFs) are depicted by solid arrows with their predicted gene products shown underneath; (**B**) Gene synteny of the region around *rokL6* in selected *Streptomyces* species. Synteny analysis was performed by Synttax (scores are given). Homologous genes are presented in the same colors. Homologs of *rokL6* and its neighbor SCO1448 are indicated in the dashed red box.

Given the importance of ROK-family regulators in the control of sugar utilization (Lu *et al.*, 2020) and the likelihood of the involvement of RokL6 in GlcN metabolism, we investigated its potential role in GlcN utilization by measuring the dry weights at different time points to compare the growth patterns of *S. coelicolor* M145 and its *rokL6* mutant in NMMP containing 1% GlcN as the sole carbon source. The mutant grew well in NMMP supplemented with either 1% glucose or with 1% GlcN (**Fig. S3**).

Transcriptome analysis of the *rokL6* mutant

To obtain further insights into the regulon of RokL6 and its relationship to GlcN metabolism, RNA sequencing (RNA-Seq) analysis was performed using RNA extracted from *S. coelicolor*

M145 and its *rokL6* mutant grown on MM agar with either mannitol or mannitol + 50 mM GlcN as the carbon sources. Biomass was harvested at two time points corresponding to vegetative growth or sporulation. In total, 29 genes were significantly differentially expressed under at least one of the conditions analyzed, using an adjusted *P*-value < 0.01 and a log2 fold change > 2 or < −2 as threshold (**Table S3**). Genes SCO1446 and SCO1448-SCO1450, which flank *rokL6*, were upregulated in the *rokL6* mutant under all tested conditions (**Fig. 3**). In addition, SCO0476 (for an unknown ABC transport protein) was downregulated at 24 h. In the absence of GlcN, transcription of the *pstSCA* operon (SCO4140–4142), which is part of the PhoP regulon and related to the transport of inorganic phosphate (Martín & Liras, 2021), was significantly upregulated (greater than fourfold), while SCO5338 (probable regulatory protein) and SCO5339 (probable plasmid transfer protein) were downregulated in Δ*rokL6* during sporulation (**Fig. 3A**). In the presence of GlcN, the putative ABC transporter genes SCO3704-SCO3706 were upregulated in Δ*rokL6* during sporulation in the presence of GlcN; transcription of genes in the biosynthetic gene cluster for carotenoid biosynthesis (Takano *et al.*, 2005), including *crtE* (SCO0185), *crtI* (SCO0186), *crtB* (SCO0187), *crtY* (SCO0191), *litQ* (SCO0192), and the regulatory gene *litR* (SCO0193), was all strongly downregulated in the *rokL6* mutant during vegetative growth (**Fig. 3B**). Indeed, compared to the wild-type M145, the production of carotenoids was significantly reduced in Δ*rokL6* in the presence of GlcN when plates were incubated in the light (**Fig. S4**), suggesting that RokL6 plays a role in the control of carotenoid biosynthesis in *S. coelicolor*.

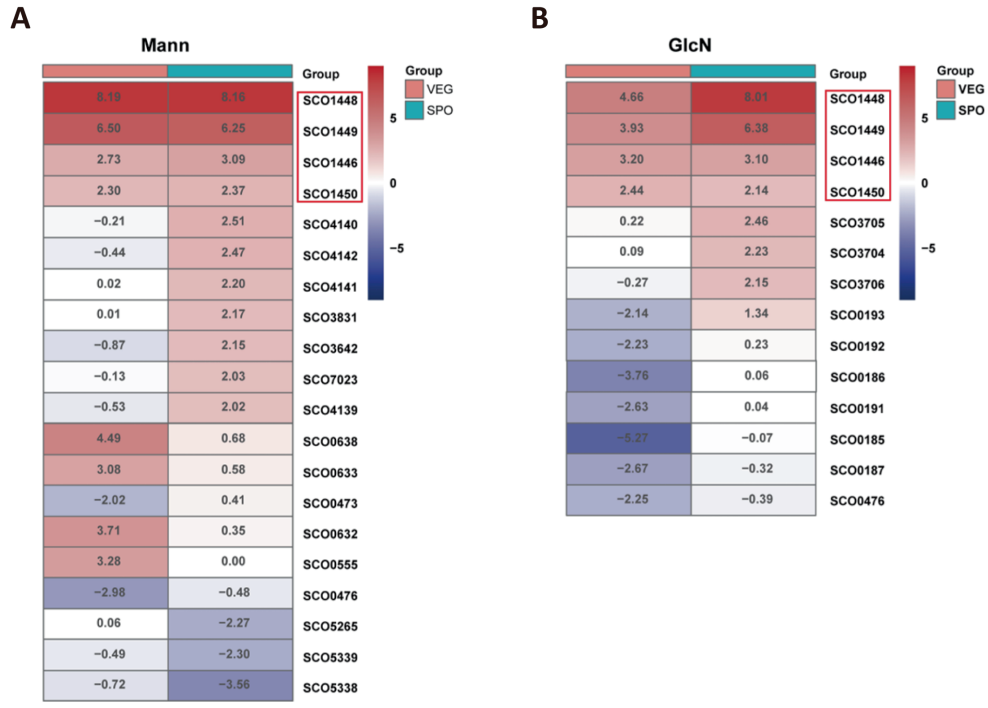

Figure 3. Heat maps of genes differentially expressed between the *rokL6* mutant and its parent *S. coelicolor* M145. Transcription patterns (expressed as log2 fold changes Δ*rokL6*/wild-type) are presented for genes differentially expressed when grown on MM with mannitol (**A**) or on MM with mannitol and GlcN (**B**). Only genes with an adjusted *p*-value less than 0.01 are shown. Navy, down-regulated (log2 fold change < -2) and brick red, up-regulated (log2 fold change > 2) in the *rokL6* mutant; intermediate log2 fold changes represented in white. Genes significantly differentially expressed under all growth conditions are highlighted with red boxes. VEG, vegetative phase; SPO, sporulation phase.

RokL6 specially binds to the intergenic region of *rokL6* and SCO1448

To identify direct targets of RokL6 *in vivo*, we performed chromatin immuno-precipitation combined with sequencing. For this, we constructed a strain expressing RokL6 with a $FLAG_3$ tag, as described in the **Materials and Methods** section. We used CRISPR-Cas9 technology to construct strain *S. coelicolor* M145-*rokL6*-FLAG, which has an engineered in-frame 3× FLAG epitope fused *in situ* and before the stop codon of *rokL6*. The functionality of RokL6-$FLAG_3$ was verified by introducing pCOM-FLAG into Δ*nagB*Δ*rokL6*. Indeed, while Δ*nagB*Δ*rokL6* mutants grow well on MM with mannitol and GlcN, Δ*nagB*Δ*rokL6*-FLAG expressing RokL6-$FLAG_3$ failed to grow (**Fig. S5**), showing that the RokL6-$FLAG_3$ was indeed expressed and active. ChIP-Seq analysis was performed with M145-*rokL6*-FLAG using cultures grown for 24 and 48 h on MM with mannitol and MM with mannitol and GlcN. Under all conditions tested, only one major binding event was observed in all of the samples, namely to the intergenic region of *rokL6* and

SCO1448 (**Fig. 4A**). Binding of RokL6 to the intergenic region shared between *rokL6* and SCO1448 was further verified *in vitro* by EMSAs. For this, C-terminally 6× His-tagged RokL6, RokL6-His_6, was expressed and purified from *E. coli* BL21(DE3). Compared to the probe for the negative control, *hrdB* promoter (*hrdB*-P), which exhibited no retardation in the presence of RokL6, the result shows that RokL6 caused a strong retardation on the probe corresponding to the promoter region of *rokL6*-SCO1448 (SCO1448-P). Specifically, almost all SCO1448-P was bound when 1 µM RokL6-His_6 was added in the EMSA (**Fig. 4B**).

Table 1. Alignment of palindromic sequences found upstream of genes encoding RokL6 and its orthologs in streptomycetes

Microorganism	Gene name	Pos*	Sequence#
S.coelicolor	*rokL6*	-64	cgtaaactatcagggagCctgcctgatagatagtggatta
Streptomyces lividans	SLIV_30530	-64	cgtaaactatcagggagCctgcctgatagatagtggatta
*Streptomyces CCM*_MD2014	NI25_31970	-65	cgtaaactatcagggagCctgcctgatagatagtgcatta
Streptomyces ambofaciens	SAM40697_1335	-64	cgtaaactatcaggaacCctgcctgatagataggggatat
Streptomyces pactum	B1H29_30360	-65	cgtaaactatcaggaacCctgcctgatagataggcgatta
Streptomyces parvulus	SPA_05455	-65	cgtaaactatcagggacCctgcctgatagatagtggatta
Streptomyces aquilus	EJC51_09795	-51	cgtaaactatcagggacCctgcctgatagatagagcgtca
Streptomyces chromofuscus	IPT68_06320	-55	ggtaaattatcagggacCctgcctgatagatagagggcgt
Streptomyces cacaoi subsp. asoensis	G9272_09180	-74	cgtaaactatcaggaacCctgcctgataaatagaggccga
Streptomyces venezuelae ATCC 21113	DEJ44_04995	-64	ggtaaactatcaggaacCctgcctgataaatagaccgcag
Streptomyces cadmiisoli	DN051_31430	-72	ggtaaactatcaggaacCctgcctgatagataggggaccc

*Position, distance of the central nucleotide (shown in uppercase) of the consensus sequence, relative to the translational start codon of the gene. #Bases within palindromic sequences are underlined.

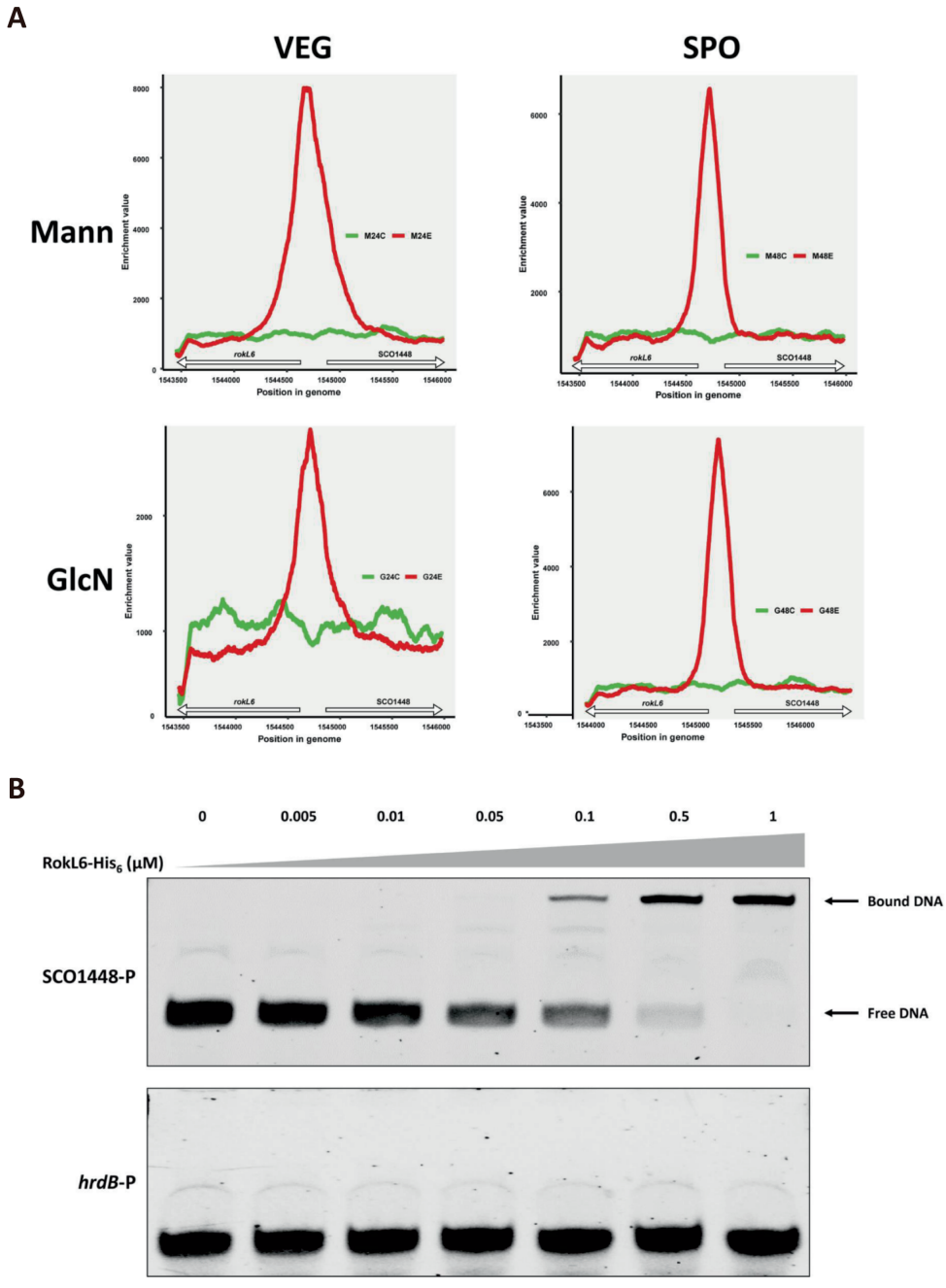

Figure 4. Identification of RokL6 binding sites *in vivo* and *in vitro*. (**A**) Enrichment peaks of RokL6 by ChIP-Seq. Red line, ChIP sample; green line, input chromosomal DNA used as a negative control. Genes flanking the peak summits are indicated. Numbers on the *X*-axis indicate genomic positions. Samples

were collected on mannitol (Mann) and mannitol with GlcN (GlcN) at two time points, vegetative phase (VEG) and sporulation phase (SPO). (**B**) EMSAs to establish direct binding *in vitro*. The binding of RokL6 to the intergenic sequence between *rokL6* and SCO1448 (SCO1448-P) was tested by EMSAs, with the *hrdB* promoter (*hrdB*-P) sequence as the negative control. Concentrations given in µM.

RokL6 binds to the inverted repeat *rokL6*-IR

Due to the high conservation in terms of both amino acid sequence and gene synteny of *rokL6* and SCO1448 with their orthologs, we reasoned that the orthologs of RokL6 may be autoregulators. To identify a putative RokL6 binding consensus, the intergenic DNA sequence of *S. coelicolor rokL6* and SCO1448, and 10 additional orthologous gene pairs from other streptomycetes (**Table 1**) were analyzed for the presence of conserved motifs using MEME. This identified a highly conserved inverted repeat sequence of 23 nucleotides in all intergenic regions (**Fig. 5A**). The consensus of the motif was C(T)TATCAGG-seven nt-CCTGATAG(A), which contains an inverted repeat designated as *rokL6*-IR, a primary candidate for the RokL6 binding site.

To verify if this is indeed a *bona fide* RokL6 binding site, EMSAs were performed using 60 bp probes. These probes contained either the intact *rokL6*-IR sequence (probe RokL6-BS) or mutant versions in which one or two regions of *rokL6*-IR were replaced by *Eco*RI or *Bam*HI restriction sites (probe RokL6-M1 to M3). Compared to RokL6-BS, which was bound at all tested concentrations (0.1, 0.2, and 0.5 µM) of RokL6-His_6, the binding signals of RokL6-His_6 to the mutated probes RokL6-M1 and RokL6-M2 were weakened, and to RokL6-M3 were fully abolished (**Fig. 5B**). These data show that the inverted repeat sequence *rokL6*-IR is indeed an essential part of the RokL6 binding site.

To determine if RokL6 has other binding sites on the *S. coelicolor* genome, the sequence of *rokL6*-IR was used to scan the *S. coelicolor* genome using the PREDetector algorithm (Hiard et al., 2007). In addition to the promoter regions of *rokL6* and SCO1448, one putative motif was identified with a relative high score upstream of SCO0137 encoding a PTS EII with an unknown substrate (**Table S4**). EMSAs showed that binding between RokL6 and the promoter of SCO0137 (SCO0137-P) in vitro was very weak (**Fig. S6A**). The putative PTS transporter EII genes SCO0137 and SCO0136 are co-expressed from a single transcriptional unit; to see if the operon would be the major GlcN transporter, a mutant was created in both the wild-type strain *S. coelicolor* M145 and in the *nagB* mutant. For this, the region from nt position +10 of SCO0137 to nt position +1,542 of SCO0136 was replaced by *aac(3)IV* (see **Materials and Methods**). The deletion of SCO0136-SCO0137 in the *nagB* mutant did not relieve GlcN toxicity (Fig. S6B). This strongly suggests that SCO0136-SCO0137 are at least not solely responsible for GlcN transport in *S. coelicolor*.

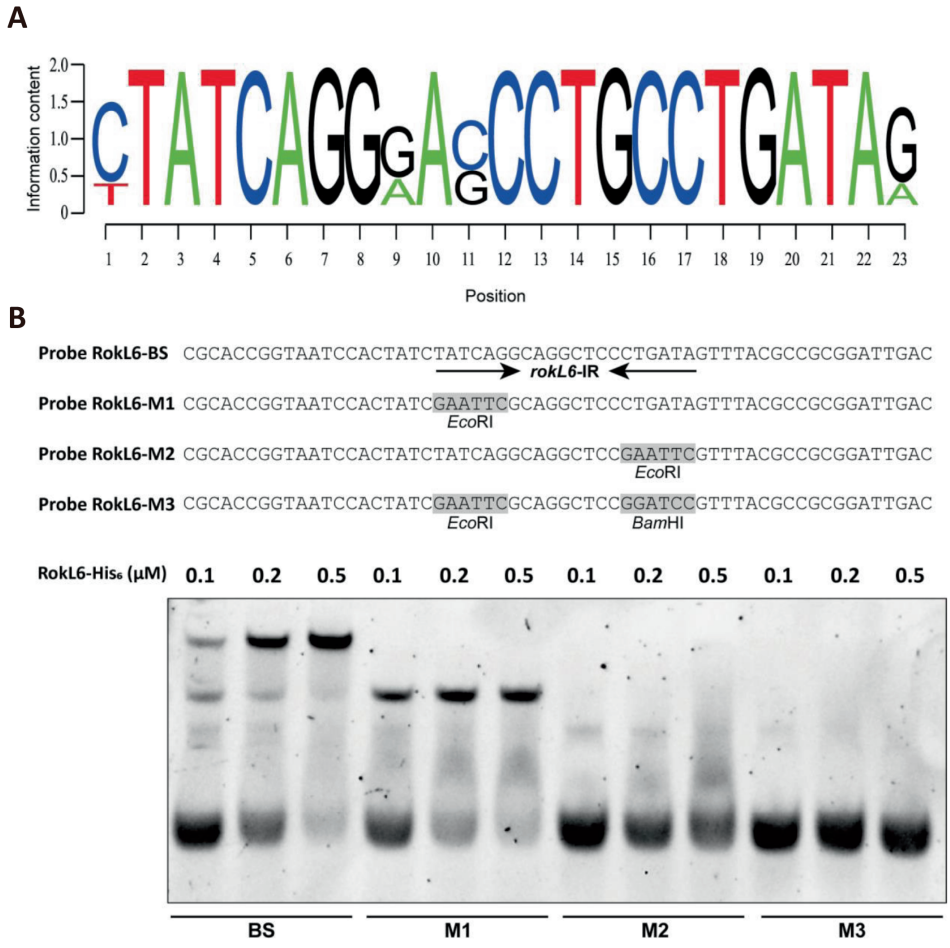

Figure 5. Identification of the RokL6 binding motif. (**A**) Sequence Logo representation of consensus sequence of RokL6 binding sites (analyzed by WebLogo). (**B**) RokL6 binding sites confirmation by EMSA. EMSAs of RokL6-His_6 with four probe sequences (RokL6-BS, RokL6-M1, RokL6-M2, RokL6-M3) were tested. The palindromic sequences (*rokL6*-IR) in Probe RokL6-BS are indicated by arrows. For each probe tested, the concentrations of RokL6-His_6 were 0.1, 0.2, and 0.5 μM.

Identification of the transcription start sites for *rokL6* and SCO1448 and repression by RokL6

5′ rapid amplification of cDNA ends (5′ RACE) was applied to identify the promoters of both *rokL6* and SCO1448. Fragments of 5′ cDNA of *rokL6* and SCO1448 were PCR-amplified from the *S. coelicolor* genome and sequenced (**Fig. S7**). The TSS of *rokL6* was localized to a G located 67 nt upstream of the *rokL6* translation start codon (TSC), and the SCO1448 TSS was mapped to a G located 27 nt upstream of the SCO1448 TSC (**Fig. 6A**). The RokL6 binding site *rokL6*-IR

encompasses nt positions −7 to +14 relative to the *rokL6* TSS, which is immediately downstream of the putative −10 region of the *rokL6* promoter. Suggestively, the *rokL6*-IR is located exactly between the −10 and −35 sequences of the SCO1448 promoter. This places the RokL6 binding site in the ideal position to repress both genes at the same time. Binding of the repressor close to the −10 or −35 consensus boxes for the RNA polymerase sigma factor is common, as it interferes with the recognition, binding of RNA polymerase, and the formation of the transcriptional initiation complex (Ruff *et al.*, 2015). This suggests that RokL6 represses the transcription of both SCO1448 and *rokL6* by binding to the core promoter regions to prevent their transcription from proceeding properly.

To determine if this is indeed the case, promoter activity assays were performed using eGFP as the reporter gene cloned in integrative vector pSET152, using promoters P*rokL6* (*rokL6* promoter) and P1448 (SCO1448 promoter), with P30 and the DasR-controlled *nagKA* promoter (P*nagKA*)as the controls. The resulting vectors pEGFP-*rokL6*, pEGFP-1448, pEGFP-*nagKA*, and pEGFP-P30 were introduced into *S. coelicolor* M145 and Δ*rokL6*. The recombinant strains were grown on MM agar with 1% mannitol and analyzed for eGFP expression levels by fluorescence microscopy and qPCR. When eGFP was expressed from either the *nagKA* promoter or the artificial promoter P30, fluorescence was unchanged between M145 and Δ*rokL6*. Conversely, when expressed from either P*rokL6* or P1448, eGFP levels were significantly higher in Δ*rokL6* as compared to the parental strain M145 (**Fig. 6B**). This result was confirmed by qPCR for the transcription of the gene for eGFP, with enhanced expression of eGFP in the *rokL6* mutant as compared to the parent, while again no significant differences in expression levels were observed for either P*nagKA* or P30. However, the transcript levels of the gene for eGFP from both P*rokL6* and P1448 were significantly upregulated in Δ*rokL6* (**Fig. 6C**), suggesting that RokL6 indeed represses the transcription of SCO1448 and at the same time acts as an autoregulator.

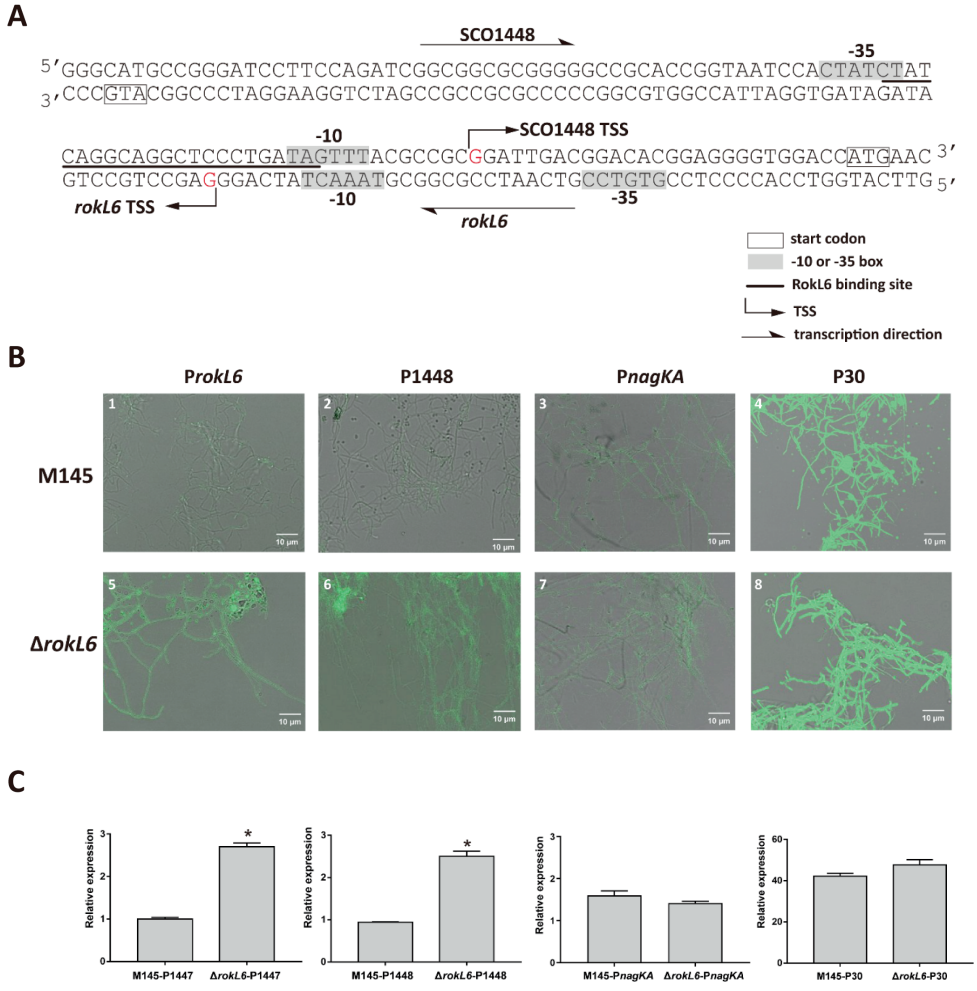

Figure 6. RokL6 acts as transcriptional repressor of SCO1448 and as auto-repressor. (**A**) Nucleotide sequences of *rokL6* and SCO1448 promoter region and RokL6-binding site. Number, distance (nt) from the respective translation start codon (TSC); bent arrows, transcription start sites (TSS); boxes, start codons; grey shading, predicted -10 or -35 boxes based on 5'RACE; underlined, RokL6 binding site; straight arrow, direction of transcription. (**B**) Fluorescence intensity of eGFP measured based on confocal fluorescence micrographs. Mycelia of the strains transformed by eGFP with different promoters were imaged using confocal microscopy. Strains analysed as indicated: 1, M145-P*rokL6;* 2, M145-P1448; 3, M145-P*nagKA*; 4, M145-P30; 5, Δ*rokL6*-P*rokL6*; 6, Δ*rokL6*-P1448; 7, Δ*rokL6*-P*nagKA*; 8, Δ*rokL6*-P30. (**C**) Relative transcription levels of eGFP. Transcription levels of for eGFP gene expressed from different promoters (P*rokL6*, P1448, P*nagKA* and P30) in M145 and Δ*rokL6* were measured by qPCR analysis. Relative transcription levels were normalized to the transcription level of the gene for eGFP with P*rokL6* in M145, which was set as 1. Data were calculated from triplicate biological experiments and presented as mean ± SD (* $P < 0.05$).

Overexpression of transporter SCO1448 relieves GlcN toxicity in *nagB* mutants

The data above demonstrate that putative transporter protein SCO1448 is the primary target of RokL6. We therefore hypothesized that SCO1448 may function as an exporter of one or more toxic metabolic intermediates that accumulate in *nagB* mutants grown on media containing GlcN. Inactivation of the repressor gene *rokL6* results in enhanced expression of SCO1448, alleviating GlcN toxicity, presumably by exporting a toxic intermediate. If this is indeed the case, expressing SCO1448 from a RokL6-independent (constitutive) promoter should have the same effect as deleting *rokL6*. To test this, SCO1448 was placed under the control of the strong and constitutive P30 promoter (see **Materials and Methods** for details) and introduced into Δ*nagB* to obtain SCO1448 overexpression strain, Δ*nagB*SCO1448-O. As a control, we used empty plasmid pSET152, to obtain Δ*nagB*-E. Transcription levels of SCO1448 were measured, showing that in both the de-repressed Δ*nagB*Δ*rokL6* and the overexpressing Δ*nagB*SCO1448-O strains, SCO1448 was highly expressed (**Fig. 7A**). Importantly, RokL6-independent expression of SCO1448 fully alleviated the toxicity of GlcN to *nagB* mutants, while Δ*nagB* or Δ*nagB* with the empty plasmid were still sensitive to GlcN (**Fig. 7B**).

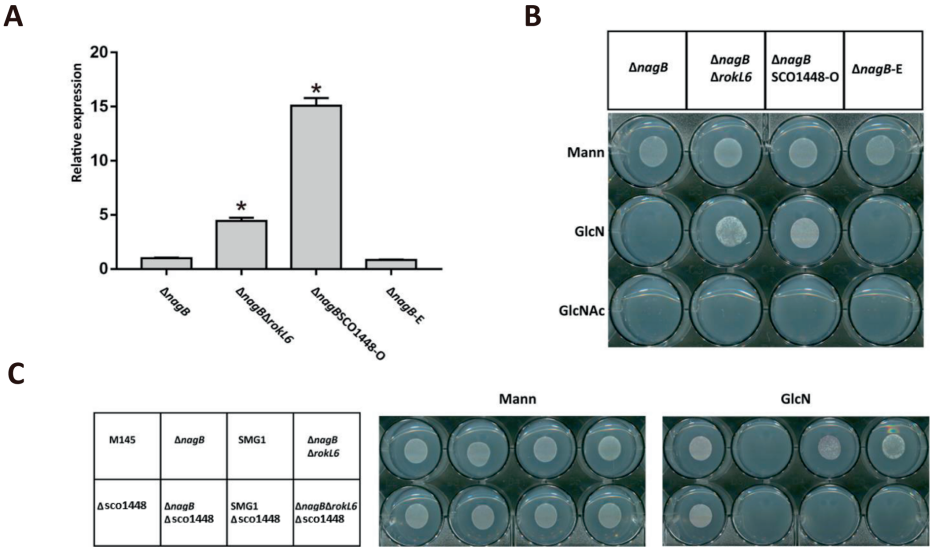

Figure 7. Constitutive expression of SCO1448 relieves toxicity of *nagB* mutants to GlcN. (**A**) SCO1448 transcription levels analysed by qPCR of the strains as follows, Δ*nagB*, Δ*nagB*Δ*rokL6*, Δ*nagB* with overexpressed SCO1448 (Δ*nagB*SCO1448-O), and Δ*nagB* complemented with empty pSET152 (Δ*nagB*-E). qPCR data were calculated from three independent experiments and presented as mean ± SD (* $P < 0.05$). (**B**) Growth of Δ*nagB*SCO1448-O on MM with GlcN(Ac). Spores of Δ*nagB*, Δ*nagB*Δ*rokL6*, Δ*nagB*SCO1448-O and Δ*nagB*-E, as indicated on the top, were spotted on MM with mannitol (Mann), MM with 1% mannitol and 5 mM GlcN (GlcN), and MM with 1% mannitol and 5 mM GlcNAc (GlcNAc) to determine their growth. (**C**) Growth of SCO1448 knock-out strains on GlcN. For media see (**B**).

Finally, we evaluated the viability of SCO1448 knock-out mutants in the presence of GlcN. For this, the coding regions of SCO1448, spanning nucleotides + 6 to +1,023 (relative to the translational start site), were substituted with apramycin resistance cassette *aac(3)IV* in *S. coelicolor* M145, Δ*nagB*, SMG1 (Δ*nagB* with a suppressor mutation in *rokL6*), and Δ*nagB*Δ*rokL6*. Homologous recombination of the gene was achieved by the introduction of vector pKO-1448, which was constructed by cloning the resistance cassette between the upstream and downstream flanking regions of SCO1448 in the unstable multi-copy plasmid pWHM3. Correct recombination events were confirmed by resistance to apramycin and sensitivity to thiostrepton (for loss of the plasmid) and by PCR, thus obtaining SCO1448 knock-out mutants, ΔSCO1448, Δ*nagB*ΔSCO1448, SMG1ΔSCO1448, and Δ*nagB*Δ*rokL6*ΔSCO1448. As expected, when SCO1448 was deleted in suppressor mutant SMG1 or in Δ*nagB*Δ*rokL6*, neither strain could grow in the presence of GlcN (**Fig. 7C**). This again provides compelling evidence that SCO1448 is solely responsible for alleviating the toxicity of GlcN to *nagB* mutants. Taken together, our data show that RokL6 directly represses the expression of SCO1448 and that, in turn, SCO1448 is responsible for alleviating GlcN toxicity to *nagB* mutants. It is likely that SCO1448 exports toxic metabolic intermediates derived from GlcN-6P, which accumulate in *nagB* mutants of *S. coelicolor* when grown on GlcN.

Discussion

The amino sugar *N*-acetylglucosamine is a preferred nutrient for *Streptomyces* and also acts as a signalling molecule for the nutritional status of the environment. While GlcNAc metabolism and transport in streptomycetes have been well studied, little is known about how GlcN is metabolized in *Streptomyces*. *S. coelicolor nagB* mutants, which cannot convert GlcN-6P into the glycolytic intermediate Fru-6P, fail to grow on either GlcNAc or GlcN, indicating that toxic intermediates are produced when GlcN-6P is not actively metabolized by NagB. We used this feature to select for suppressor mutants, whereby it is important to note that some of the *nagB* suppressors selected on GlcNAc fail to confer resistance to GlcN and vice versa. In this work, we focused on the GlcN-specific gene *rokL6* (SCO1447), the mutation of which relieves GlcN toxicity in *nagB* mutants.

RokL6 is an ROK-family regulator, and members of this family of regulators often play a role in the control of sugar metabolism. Systems-wide analysis using RNA-Seq, qPCR, ChIP-Seq, and EMSAs showed that RokL6 directly represses the transcription of SCO1448, which is annotated as an MFS sugar transporter. SCO1448 plays a key role in alleviating GlcN toxicity in the absence of NagB activity. Upregulation of SCO1448 is sufficient to relieve GlcN toxicity to *S. coelicolor nagB* mutants, while conversely, deletion of SCO1448 in *S. coelicolor* Δ*nagB*Δ*rokL6* or the suppressor mutant SMG1 abolishes the acquired GlcN resistance (**Fig. 7**). Thus, the key to GlcN resistance lies in the expression of SCO1448. Still, a lot is unclear about the *rokL6*-

SCO1448 gene pair. In particular, why would streptomycetes have a cryptic exporter to protect themselves from toxic intermediates related specifically to GlcN? The system is likely important, as its gene synteny and also (the location of) the RokL6 binding site are highly conserved in *Streptomyces* (**Fig. 2B** and **Table 1**). The high conservation of the binding site predicts that the expression of SCO1448 is repressed by RokL6 in many if not all streptomycetes. While the biological significance is not yet understood, the proteins likely play an important role in preventing the accumulation of excess toxic intermediates under specific growth conditions in the natural environment. How the repression by RokL6 is relieved in the cell and what the possible ligands are that control its DNA binding, which likely facilitates the derepression of SCO1448, remains to be elucidated.

An important question is also what is the exact nature of the toxic molecule(s) that accumulate(s) in *nagB* mutants, which are likely derived from GlcN-6P. Mutants that lack a functional NagA enzyme accumulate high levels of GlcNAc-6P, which is lethal in *E. coli* and *Bacillus subtilis* (Plumbridge, 2009, Plumbridge, 2015, White, 1968). However, *S. coelicolor nagA* null mutants are able to grow in the presence of either GlcNAc or GlcN, and deletion of *nagA* from *nagB* null mutants of *S. coelicolor* also alleviates the toxicity of either amino sugar (Świątek *et al.*, 2012b). Still, there are no metabolic routes that point to a key role for NagA in GlcN metabolism. Why then would mutation of *nagA* prevent toxicity of not only GlcNAc but also GlcN in *nagB* mutants? This again shows that there are still significant gaps in our understanding of amino sugar metabolism in *Streptomyces*. This is in fact quite surprising for such an important central metabolic pathway.

ChIP-Seq and EMSA assays demonstrated that RokL6 binds directly to the palindromic *rokL6*-IR identified from the intergenic region of *rokL6* and SCO1448, with the sequence 5'-C(T)TATCAGG-seven nt-CCTGATAG(A)-3'. Furthermore, the precise transcription start sites for *rokL6* and SCO1448 were identified by 5' RACE, showing that the two genes are transcribed from overlapping promoters. RokL6 strategically binds to a site that is located precisely in between the −35 and −10 boxes for SCO1448 and downstream of the −10 box for *rokL6*, allowing RokL6 to inhibit the transcription of both genes at the same time. The transcriptional repression of SCO1448 and autoregulation of *rokL6* by RokL6 were further demonstrated by promoter activity assays and qPCR (**Fig. 6**), showing that RokL6 indeed inhibits the transcription of both genes.

The RokL6 binding site is distinct from operators characterized for ROK-family regulators in, e.g., *E. coli* and Firmicutes. The operator consensus sequences of NagC and Mlc in *E. coli* or XylR in firmicutes are typically composed of two A/T-rich inverted repeats separated by a spacer of 5–9 bp (Dubeau *et al.*, 2011, El Qaidi & Plumbridge, 2008, Gu *et al.*, 2010). The binding consensus sequences of RokB from *S. coelicolor* and RokA from *Streptococcus pneumoniae* are also enriched with T and A at the 5'-end and 3'-end (Shafeeq *et al.*, 2012, Bekiesch *et al.*, 2016), similar to CysR from *Corynebacterium glutamicum* (Rückert *et*

al., 2008). The binding target of CsnR from *Streptomyces lividans* shares some similarity with that of RokL6, containing the sequence 5'-AGG-seven nt-CCT-3' in the binding consensus. In contrast, the binding consensus of Rok7B7 of *Streptomyces avermitilis* is 5'-TTKAMKHSTTSAV-3' and is unrelated to that of RokL6 (Dubeau *et al.*, 2011, Lu *et al.*, 2020).ChIP-Seq experiments identified a single binding site in all samples. Using that binding consensus as input, scanning the *S. coelicolor* genome using PREDetector revealed one additional sequence with similarity to the RokL6 consensus, namely upstream of the SCO0137-SCO0136 operon that encodes PTS transporter EIIC enzymes. Its control by RokL6 and the fact that in *E. coli* and *B. subtilis*, GlcN is transported via the PTS (Plumbridge, 2000, Zeppenfeld *et al.*, 2000, Gaugué *et al.*, 2013), suggested a possible role for the SCO0137-SCO0136 operon in GlcN transport. We tested whether the inactivation of the operon would affect GlcN toxicity in *nagB* mutants, but this is not the case (**Fig. S6B**). Therefore, PTS transporter SCO0136-SCO0137 is at least not the only transporter for GlcN uptake in *S. coelicolor*, and more studies are required to understand how GlcN is internalized in streptomycetes.

In terms of indirect effects of the deletion of *rokL6*, besides SCO1448, also the adjacent SCO1446 and SCO1448-SCO1450 were upregulated in *rokL6* mutants. Other genes whose transcription was affected were those for transporters SCO0476, SCO3704-SCO3706, and SCO4140-SCO4142, as well as genes of the carotenoid biosynthetic gene cluster (BGC). Upregulation of the latter BGC was validated by enhanced pigmentation of mycelia grown in the light (**Fig. S4**). However, our experiments strongly suggest that the intergenic region between *rokL6* and SCO1448 is the primary binding site for RokL6, and the other genes whose transcription was affected in *rokL6* mutants are likely controlled indirectly.

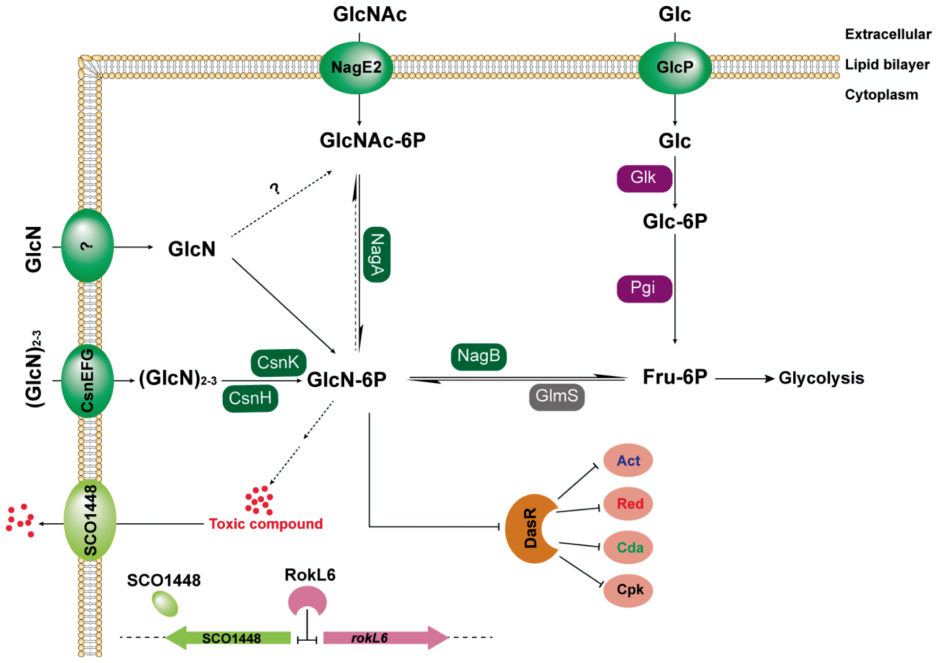

Figure 8. Model for amino sugar metabolism and the roles of RokL6 and SCO1448 in *S. coelicolor*. Metabolism of GlcNAc is well understood in *S. coelicolor*. GlcN-6P is an allosteric inhibitor of the global repressor DasR, which represses the biosynthesis of antibiotics. Based on our data, we propose that toxic substances are formed from GlcN metabolism, in particular in *nagB* mutants; these are likely derived from GlcN-6P or its derivatives, and are exported by the MFS transporter SCO1448. We show that this transporter is transcriptionally repressed by the ROK-family regulator RokL6, and that expression of SCO1448 relieves GlcN toxicity. Metabolic routes are presented by black arrows with the enzymes as indicated/unknown routes by dotted arrows. Unknown transporters and enzymes are indicated by question marks. Note that toxic intermediates arising from GlcNAc are not relieved by SCO1448 and hence require at least in part another metabolic route. For details, see the main text. For abbreviations not mentioned in the text: Glc, glucose; Glc-6P, glucose-6-phosphate; GlcP, glucose permease, Glk, glucokinase; Pgi, glucose-6-phosphate isomerase; Glms, glucosamine-fructose-6-phosphate aminotransferase; CoA-Ac, Acetyl Coenzyme A; Act, actinorhodin; Red, prodiginines; Cda, calcium-dependent antibiotic; Cpk, cryptic polyketide.

Our current understanding of amino sugar metabolism and transport in *Streptomyces* is shown schematically in **Fig. 8**, including possible functions for RokL6 and SCO1448. GlcN-6P is derived either from the internalization of GlcN or the deacetylation of GlcNAc-6P. High accumulation of GlcN-6P and/or their metabolic derivatives is lethal to *S. coelicolor*, in line with observations in *E. coli* and *B. subtilis*. The MFS transporter SCO1448 likely serves to facilitate

the export of toxic intermediate(s), and in turn, the transcription of SCO1448 is repressed by RokL6 (**Fig. 8**). Constitutive expression of SCO1448 does not relieve the toxicity of GlcNAc, suggesting that GlcNAc and GlcN toxicity are not caused by the same molecule(s).

Taken together, this study characterizes the function of the ROK-family regulator RokL6, including its regulons, binding sites, and its roles in the relief of GlcN toxicity in *S. coelicolor*. These findings shed new light on amino sugar sensitivity and on the control of amino sugar metabolism in *Streptomyces*, which plays a central role in their life cycle.

Acknowledgments

We are very grateful to Jackie Plumbridge for stimulating discussions on amino sugar toxicity and to Ellen de Waal for technical assistance. This work was supported by a fellowship (grant number 201904910552) from the Chinese Scholarship Council (CSC) to C.L. and by the Advanced Grant Committee (grant number 101055020) of the European Research Council to G.P.v.W.

Supplementary Information for Chapter 3

Supplementary tables

Table S1. Bacterial strains and plasmids.

Bacterial strains	Descrip1on[a,b]	Reference or source
S. coelicolor A3(2)		
M145	*S. coelicolor* A3(2) M145, SCP1$^-$ SCP2$^-$	(Kieser *et al.*, 2000)
Δ*nagB*	M145 *nagB*^{d}	(Świątek *et al.*, 2012a)
SMG1	Δ*nagB* suppressor mutant	(Swiatek, 2012)
Δ*nagB*Δ*rokL6*	M145 *nagB*^{d} *rokL6*^{d}	This study
Δ*nagB*Δ*rokL6*C	Δ*nagB*Δ*rokL6* complemented with *rokL6*	This study
Δ*nagB*Δ*rokL6*E	Δ*nagB*Δ*rokL6* conjugated with empty pSET152	This study
Δ*nagB*-pSET	Δ*nagB* conjugated with empty pSET152	This study
Δ*nagB*SCO1448OE	Δ*nagB* with overexpressed SCO1448	This study
M145-*rokL6*-FLAG	3×FLAG fused to RokL6 C-terminator in M145	This study
Δ*nagB*Δ*rokL6*-FLAG	Δ*nagB*Δ*rokL6* complemented with pCOM-$FLAG_3$	This study
ΔSCO1448	M145 SCO1448^d	This study
Δ*nagB*ΔSCO1448	Δ*nagB* SCO1448^d	This study
SMG1ΔSCO1448	SMG1 SCO1448:: *aac(3)IV*	This study
Δ*nagB*Δ*rokL6*ΔSCO1448	Δ*nagB rokL6*^{d} SCO1448:: *aac(3)IV*	This study
M145-P*rokL6*-eGFP	M145 with eGFP with the promoter P*rokL6*	This study
M145-P1448-eGFP	M145 with eGFP with the promoter P1448	This study
M145-P*nagKA*-eGFP	M145 with eGFP with the promoter P*nagKA*	This study
M145-P30-eGFP	M145 with eGFP with the promoter P30	This study
Δ*rokL6*-P*rokL6*-eGFP	Δ*rokL6* with eGFP with the promoter P*rokL6*	This study
Δ*rokL6*-P1448-eGFP	Δ*rokL6* with eGFP with the promoter P1448	This study
Δ*rokL6*-P*nagKA*-eGFP	Δ*rokL6* with eGFP with the promoter P*nagKA*	This study
Δ*rokL6*-P30-eGFP	Δ*rokL6* with eGFP with the promoter P30	This study
M145ΔSCO0136-0137	M145 SCO0136-0137:: *aac(3)IV*	This study
Δ*nagB*ΔSCO0136-0137	Δ*nagB* SCO0136-0137:: *aac(3)IV*	This study
E. coli		
E. coli JM109	General cloning strain	(Sambrook *et al.*, 1989)
E. coli ET12567/pUZ8002	Strain used for conjugation between *E. coli* and *Streptomyces*	(MacNeil *et al.*, 1992)
E. coli BL21 (DE3)	Strain used for protein expression	Novagen

Plasmids	Descrip1on	Reference or source
pWHM3	*E. coli*/*Streptomyces* shuttle vector, high copy number and unstable in *Streptomyces*	(Vara *et al.*, 1989)
pUWLcre	*E. coli*/*Streptomyces* shuttle vector expressing the Cre recombinase in *Streptomyces*	(Fedoryshyn *et al.*, 2008)
pSET152	Integrative *E. coli*/*Streptomyces* shuttle vector	(Bierman *et al.*, 1992)
pET28a (+)	Vector for His_6-tagged protein expression	Novagen

pCRISPomyces-2	*E. coli*/*Streptomyces* shuttle vector, harbouring codon optimised cas9, designed for easy inserted spacer of specific genes	(Cobb *et al.*, 2015)
pKO-*rokL6*	*rokL6* knock out recombinant vector	This study
pKO-1448	SCO1448 knock out recombinant vector	This study
pCOM-*rokL6*	*rokL6* complemented vector based on pSET152	This study
pOE-1448	SCO1448 overexpressed vector based on pSET152	This study
pKI-$FLAG_3$	3×FLAG knock in vector based on pCRISPomyces-2	This study
pCOM-$FLAG_3$	RokL6-3×FLAG complemented vector	This study
pEX-RokL6	RokL6-His_6 expressing vector based on pET28a (+)	This study
pEGFP-*rokL6*	eGFP expressing vector with *rokL6* promoter	This study
pEGFP-1448	eGFP expressing vector with SCO1448 promoter	This study
pEGFP-*nagKA*	eGFP expressing vector with *nagKA* promoter	This study
pEGFP-P30	eGFP expressing vector with P30 promoter	This study

[a]"d" indicates the gene before "d" is in-frame deleted
[b]"::*aac(3)IV*" indicates the gene before "::"is replaced by *aac(3)IV* cassette

Table S2. Oligonucleotides used un this study

Name	**5'-3' sequence**	**Func1on or descrip1ons**
Cloning used		
rokL6-KO-LF	GTCAGAATTCACCCTCGCGAACACCACCAGGCA	*rokL6* knock out
rokL6-KO-LR	GAAGTTATCCATCACCTCTAGACATGCCGGGATCCTTCCAGAT	*rokL6* knock out
rokL6-KO-RF	GAAGTTATCGCGCATCTCTAGATTCGCACCGCCGGAGCGGTAG	*rokL6* knock out
rokL6-KO-RR	GTCAAAGCTTGCATGCGCAGGCCGTCAAGC	*rokL6* knock out
rokL6-COM-F	CTAGGATATCCATCGCCACGTCCGACA	RokL6 complementation
rokL6-COM-R	CATGTCTAGAGGTGAGGCCCTTCCGGG	RokL6 complementation
rokL6-KO-checkF	TGCTGCTGCCGACGGTACTC	*rokL6* mutant check
rokL6-KO-checkF	CTATCAGGGAGCCTGCCTGATAG	*rokL6* mutant check
SCO1448-KO-LF	GTCAGAATTCTGCTGCCGACGGTACTCGGGTGG	SCO1448 knock out
SCO1448-KO-LR	GAAGTTATCCATCACCTCTAGATGTGTTCATGGTCCACCCCTC	SCO1448 knock out
SCO1448-KO-RF	GAAGTTATCGCGCATCTCTAGACCGGCAGTCCTGAACGCCTCGC	SCO1448 knock out
SCO1448-KO-RR	GTCAAAGCTTTCTCCGCGATCAGGGCGATGACG	SCO1448 knock out
SCO1448-KO-checkF	CGGGCATGCCGGGATCCTTC	SCO1448 mutant check

SCO1448-KO-checkR	GGGTGCTGGTCCGGCTGGAC	SCO1448 mutant check
SCO1448-OE-F	CTAGGATATCTGTTCACATTCGAACCGTCTCTGCTTTGACATC GTGTGGCGCTTGGGTGTAAAGTCGTGGCCATCTAAGTAAGG AGTGTCCATATGAACACAGGG ACCTACGGG	SCO1448 overexpression with SP30R15
SCO1448-OE-R	GTCGACTCTAGAGTGGACTCAC	SCO1448 overexpression
rokL6-spacer-F	ACGCGGCGCCGCGGGCTACCGCTC	*rokL6* spacer assembly
rokL6-spacer-R	AAACGAGCGGTAGCCCGCGGCGCC	*rokL6* spacer assembly
rokL6-flag-LF	CAGCTATGACCATGATTACGCTCGCCCTTCATCTGCTCCAGC	*rokL6* with 3×FLAG PCR
rokL6-flag-LR	GGCGCCGCGGGCTACTTGTCGTCGTCGTCCTTGTAGTCGAT GTCGTGGTCCTTGTAGTCGCCGTCGTGGTCCTTGTAGTCCCG CTCCGGCGGTGCGAAG	*rokL6* with 3×FLAG PCR
rokL6-flag-RF	GACAAGTAGCCCGCGGCGCCCGGAAGGGCCTCACCCGG	*rokL6* CRISPR template assembly
rokL6-flag-RR	GTTGTAAAACGACGGCCAGTGCCACGCAGCGCGGCGGGG AC	*rokL6* CRISPR template assembly
RokL6-flag-checkF	CGGGCTACTTGTCGTCGT	RokL6-FLAG check
RokL6-flag-checkR	TCGACCCTGGCTGCCTGGTG	RokL6-FLAG check
RokL6-EXP-F	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATGCCC GCATCACCGAGCACC	RokL6-His_6 expression
RokL6-EXP-R	GACAGCAAATGGGTCGCGGATCCATGCCCGCATCACCGAGC AC	RokL6-His_6 expression
egfp-F	ATGGTGAGCAAGGGCGAGGAG	eGFP gene PCR
egfp-R egfp-R2	CTTGGGCTGCAGGTCGACTTTACTTGTACAGCTCGTCCATGC	eGFP gene PCR
P1447-egfp-F	GCTATGACATGATTACGAATTCGATGAGCGCCAGCCCCAACT GGC	*rokL6* promoter PCR
P1447-egfp-R	TCGCCCTTGCTCACCATGCCGGGATCCTTCCAGATCGG	*rokL6* promoter PCR
P1448-egfp-F	GCTATGACATGATTACGAATTCGATGGCGGTGAGGCGTTCG ACGAG	SCO1448 promoter PCR
P1448-egfp-R	TCGCCCTTGCTCACCATGGTCCACCCCTCCGTGTCCG	SCO1448 promoter PCR
P*nagKA*-egfp-F	TATGACATGATTACGAATTCGATCGCCGAGGACCGCCGTCAT GC	*nagKA* promoter PCR
P*nagKA*-egfp-R	CTCGCCCTTGCTCACCATCCCGGTGCCGCCCACATCGAG	*nagKA* promoter PCR
SP30R15-egfp-F	GCTATGACATGATTACGAATTCGATTGTTCACATTCGAACCGT CTC	SP30R15 PCR
SP30R15-egfp-R	TCGCCCTTGCTCACCATATGGACACTCCTTACTTAGATGG	SP30R15 PCR
SCO0136-0137-KO-LF	CAGCTATGACCATGATTACTCGTGACGGTGAAGGTTCAC	SCO0136-0137 knock out
SCO0136-0137-KO-LR	ACGAAGTTATCGCGCATCTTGCGCTCATGACAGTGGAAC	SCO0136-0137 knock out

SCO0136-0137-KO-RF	GCTATACGAAGTTATCCATCACCTGCTCCCTGAGCGGCCGTTGAC	SCO0136-0137 knock out
SCO0136-0137-KO-RR	GTTGTAAAACGACGGCCAGTGGTACGGACCGATCATGGACC	SCO0136-0137 knock out
SCO0136-37-KO-checkF	GCTGGAACCGACCGGCTTAC	SCO0136-0137 mutant check
SCO0136-37-KO-checkR	CGGCTGTCCGCCGAGATTCG	SCO0136-0137 mutant check
qPCR used		
hrdB-qPCR-F	GCACATGGTCGAGGTCATCA	*hrdB* qPCR
hrdB-qPCR-R	GGTCATGTCGAGCTCCTTGG	*hrdB* qPCR
egfp-qPCR-F	CCTCGTGACCACCCTGACCTAC	eGFP gene qPCR
egfp-qPCR-R	TTGCCGTCGTCCTTGAAGAAGATG	eGFP gene qPCR
SCO1448-qPCR-F	CCAACCTGCTCACCCTGTG	SCO1448 qPCR
SCO1448-qPCR-R	TCCAGACGGGTCTCGATCTC	SCO1448 qPCR
5'RACE used		
oligo(dT) anchor primer	GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTV	5' RACE anchor primer
anchor primer	GACCACGCGTATCGATGTCGAC	5' RACE anchor primer
rokL6-GSP	GTCGATGGGCACGGACG	*rokL6* specific primer
rokL6-NT1	ACGCCTTCGGTGCGCACGTC	*rokL6* nested primer1
rokL6-NT2	GCTTCAACTGCCCTGCCGTC	*rokL6* nested primer2
SCO1448-GSP	GCGAGGCACAGGGTGAGCAG	SCO1448 specific primer
SCO1448-GSP	GCGGCATGGCGAGGGACG	SCO1448 nested primer1
SCO1448-GSP	CTCACGCCGGCGTGGTCCTG	SCO1448 nested primer2
EMSA used		
EMSA- *rokL6*-P-F	CGCACCGGTAATCCACTATCTATCAGGCAGGCTCCCTGATAGTTTACGCCGCGGATTGAC	RokL6 BS probe annealing
EMSA- *rokL6*-P-R	GTCAATCCGCGGCGTAAACTATCAGGGAGCCTGCCTGATAGATAGTGGATTACCGGTGCG	RokL6 BS probe annealing
EMSA- *hrdB*-P-F	CGGCCCGCCCGCACCGTCGGCCCATTCCCAAGCCGGTGGTCGGCCCCTGTCCGCCGTGGA	RokL6 EMSA *hrdB*-P annealing
EMSA- *hrdB*-P-R	TCCACGGCGGACAGGGGCCGACCACCGGCTTGGGAATGGGCCGACGGTGCGGGCGGGCCG	RokL6 EMSA *hrdB*-P annealing
EMSA- *rokL6*-MP1-F	CGCACCGGTAATCCACTATCGAATTCGCAGGCTCCCTGATAGTTTACGCCGCGGATTGAC	RokL6 EMSA M1 probe annealing
EMSA- *rokL6*-MP1-R	GTCAATCCGCGGCGTAAACTATCAGGGAGCCTGCGAATTCGATAGTGGATTACCGGTGCG	RokL6 EMSA M1 probe annealing
EMSA- *rokL6*-MP2-F	CGCACCGGTAATCCACTATCTATCAGGCAGGCTCCGAATTCGTTTACGCCGCGGATTGAC	RokL6 EMSA M2 probe annealing
EMSA- *rokL6*-MP2-R	GTCAATCCGCGGCGTAAACGAATTCGGAGCCTGCCTGATAGATAGTGGATTACCGGTGCG	RokL6 EMSA M2 probe annealing

EMSA- *rokL6*-MP3-F	CGCACCGGTAATCCACTATCGAATTCGCAGGCTCCGGATCCGTTTACGCCGCGGATTGAC	RokL6 EMSA M3 probe annealing
EMSA- *rokL6*-MP3-R	GTCAATCCGCGGCGTAAACGGATCCGGAGCCTGCGAATTCGATAGTGGATTACCGGTGCG	RokL6 EMSA M3 probe annealing

Table S3. Genes differentially expressed between the *rokL6* null mutant and its parent *S. coelicolor* M145 (wt).

Gene	Annotation	log2 fold change (Δ*rokL6*/wt)[a]			
		VEG-Mann	SPO-Mann	VEG-GlcN	SPO-GlcN
SCO0555	Membrane-bound oxidoreductase	3.280	0.004	-0.051	0.210
SCO0632	RNA polymerase sigma factor SigK	3.709	0.349	-0.917	0.214
SCO0633	Hypothetical protein	3.079	0.578	-0.668	0.178
SCO0638	Lipoprotein	4.490	0.683	-0.184	0.8109
SCO1446	Hypothetical protein	2.73	3.093	3.199	3.099
SCO1448	Transporter	8.186	8.158	4.656	8.008
SCO1449	Hypothetical protein	6.500	6.247	3.926	6.382
SCO1450	Uracyl permease	2.300	2.372	2.444	2.135
SCO0473	Solute-binding lipoprotein	-2.0157	0.4071	-0.733	-0.130
SCO0476	ABC transporter ATP-binding protein	-2.981	-0.475	-2.250	-0.392
SCO3642	Hypothetical protein	-0.875	2.147	-0.125	-0.361
SCO3831	Dehydrogenase	0.012	2.167	-0.177	0.842
SCO4139	ATP-binding protein	-0.530	2.0167	0.811	-0.115
SCO4140	Phosphate ABC transporter permease	-0.212	2.510	0.440	-0.180
SCO4141	Phosphate ABC transporter permease	0.0242	2.199	0.715	-0.039
SCO4142	Phosphate-binding protein	-0.445	2.474	0.741	0.128
SCO7023	Hypothetical protein	-0.126	2.033	-0.096	0.454
SCO5265	Hypothetical protein	0.056	-2.274	-0.029	0.056
SCO5338	Regulatory protein	-0.723	-3.560	0.195	0.0240
SCO5339	Plasmid transfer protein	-0.486	-2.302	0.476	-0.078
SCO0185	Geranylgeranyl pyrophosphate synthase	0.685	0.555	-0.500	0.273
SCO0186	Phytoene dehydrogenase	0.322	0.329	0.802	0.557
SCO0187	Phytoene synthase	0.580	0.431	-2.672	-0.315
SCO0191	Lycopene cyclase	0.0560	0.424	-2.626	0.039
SCO0192	Oxidoreductase	0.194	0.459	-2.232	0.232
SCO0193	DNA-binding regulator	0.303	1.144	-2.137	1.342
SCO3706	ABC transporter ATP-binding protein	-0.117	0.355	-0.2690	2.149
SCO3705	ABC transporter membrane subunit	0.039	0.431	0.218	2.461

SCO3704	Substrate-binding transport protein	-0.203	0.140	0.0916	2.229

[a] Differentially expressed genes at least one condition between the Δ*rokL6* and wt; sample collected from MM + mannitol at 24 h (VEG-Mann), MM + mannitol at 48 h (SPO-Mann), MM + mannitol and GlcN at 24 h (VEG-GlcN), MM + mannitol and GlcN at 48 h (SPO-GlcN).

Table S4. Putative targets of RokL6 as predicted by PREDetector

Genes	Annotations	Scores*	Sequences#
SCO1448	Probable transport protein	29.4	CTATCAGGCAGGCTCCCTGATAG
rokL6	ROK family regulator	29.4	CTATCAGGCAGGCTCCCTGATAG
SCO0137	Transmembrane transport protein	15.4	CTTTCAGACATGGTTCCTGATTG
SCO4114	Sporulation associated protein	7.1	TTAAGAGGCATTGTTCCGGATGG
SCO1359	Possible integral membrane protein	6.7	CTACGAGGACGACTGCCTCATCG
SCO1358	LysR family transcriptional regulator	6.7	CTACGAGGACGACTGCCTCATCG
SCO5864	Hypothetical protein	6.2	TGATCATGCACCCTGTCGGAAAG
SCO3250	Probable integrase	6.0	CCATCAGGCAGCCTCCTCGATCT

*Sections of sequences underlined are those that match the proposed RokL6 consensus sequence.
#Matrix scores determined by PREDetector; score above 10 are considered relevant

Supplementary figures

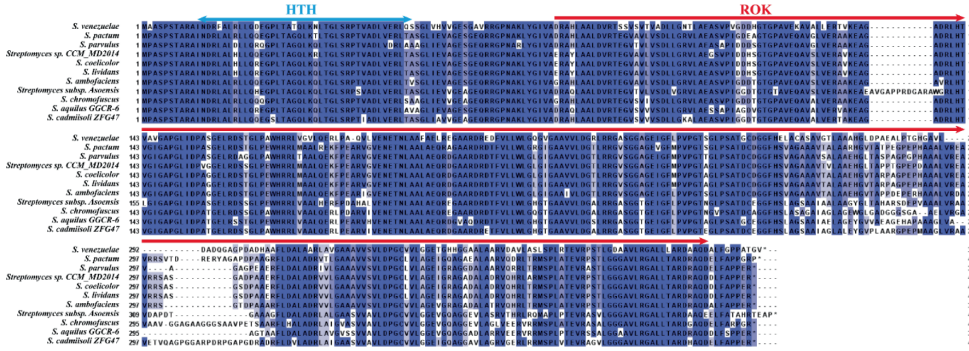

Figure S1. Alignment of RokL6 and orthologs in Streptomyces. Multiple sequence alignment was performed using the amino acid sequence of RokL6 and its orthologs from other *Streptomyces* species by Clustal Omega. Identical amino acids are shown in dark blue, and amino acids with similar properties in light blue. The helix-turn-helix (HTH) domain and ROK-kinase domain are indicated. The image was generated using Jalview (version 2.11.2.7).

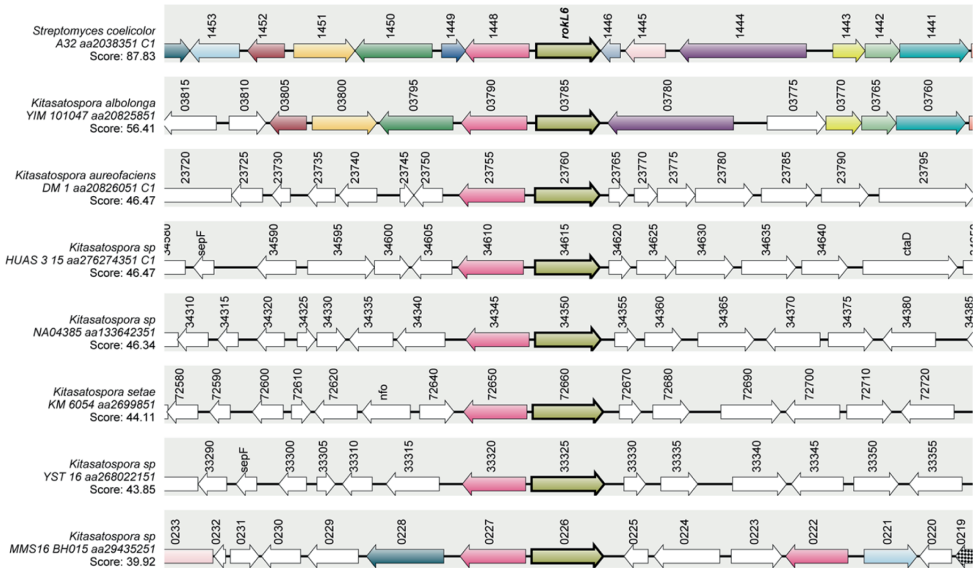

Figure S2. Gene synteny analysis of rokL6-SCO1448 in *Kitasatospora* species. Synteny analysis was performed by Synttax (scores are given). Homologous genes are presented as indicated and in the same colours.

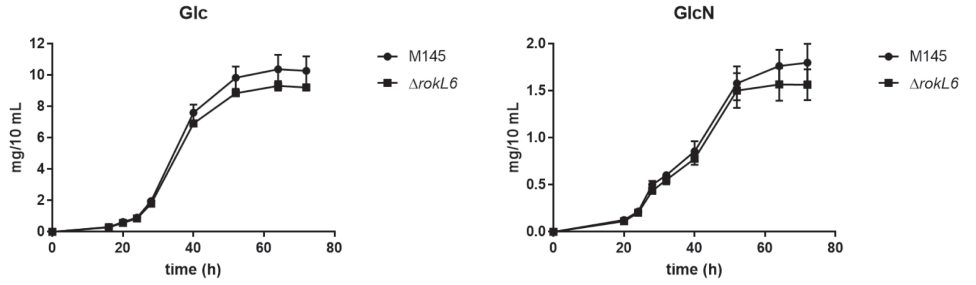

Figure S3. Growth curves of M145 and ΔrokL6 in NMMP. Comparison of the growth pattern between wild-type *S. coelicolor* M145 and the mutant Δ*rokL6*. Strains were growth in liquid NMMP supplemented with either 1% (w/v) glucose (Glc) or glucosamine (GlcN). The dry weights of the mycelia were measured at different time points. Each value was obtained from three biological replicates as a mean ± SD.

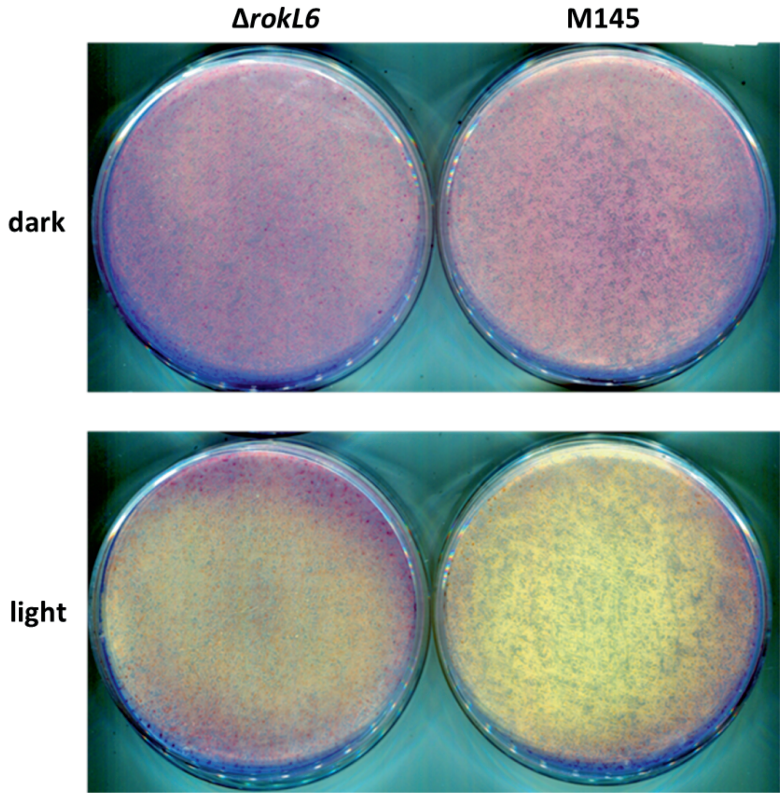

Figure S4. RokL6 up-regulates produc1on of carotenoids. 10^7 spores of wild-type M145 and its *rokL6* mutant were plated on MM with 1% mannitol and 50 mM GlcN, and cultured at 30°C for 48 h in the dark (dark) and light (light). Note that the petri dish plated with wild-type M145 appeared significantly yellower than that with Δ*rokL6* due to the upregulated production of carotenoid under the light condition (no significant difference observed in the dark).

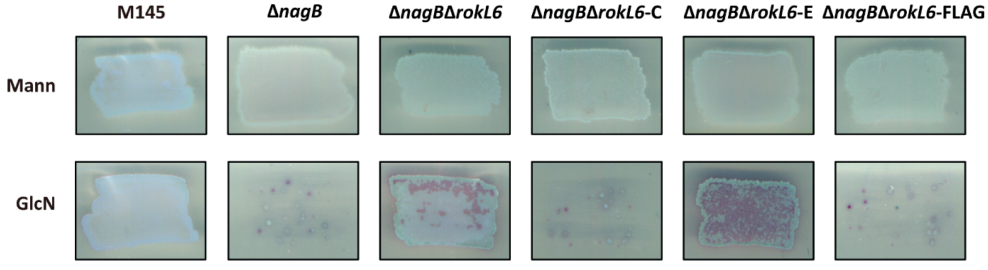

Figure S5. Verification of the functionalily of RokL6-FLAG$_3$. Spores of *S. coelicolor* M145, its mutants Δ*nagB*, Δ*nagB*Δ*rokL6*, Δ*nagB*Δ*rokL6*-C (Δ*nagB*Δ*rokL6* complemented with RokL6), Δ*nagB*Δ*rokL6*-FLAG (Δ*nagB*Δ*rokL6* complemented with RokL6-FLAG$_3$) and Δ*nagB*Δ*rokL6* harbouring empty plasmid pSET152 (Δ*nagB*Δ*rokL6*-E) were plated on MM with either mannitol (Mann) or mannitol with GlcN (GlcN), and cultured at 30°C to check forr growth on GlcN. Note that introduction of a plasmid

expressing either wild-type RokL6 (column 4) or RokL6-$FLAG_3$ (column 6) made the Δ*nagB*Δ*rokL6* double mutant sensitive to GlcN once more.

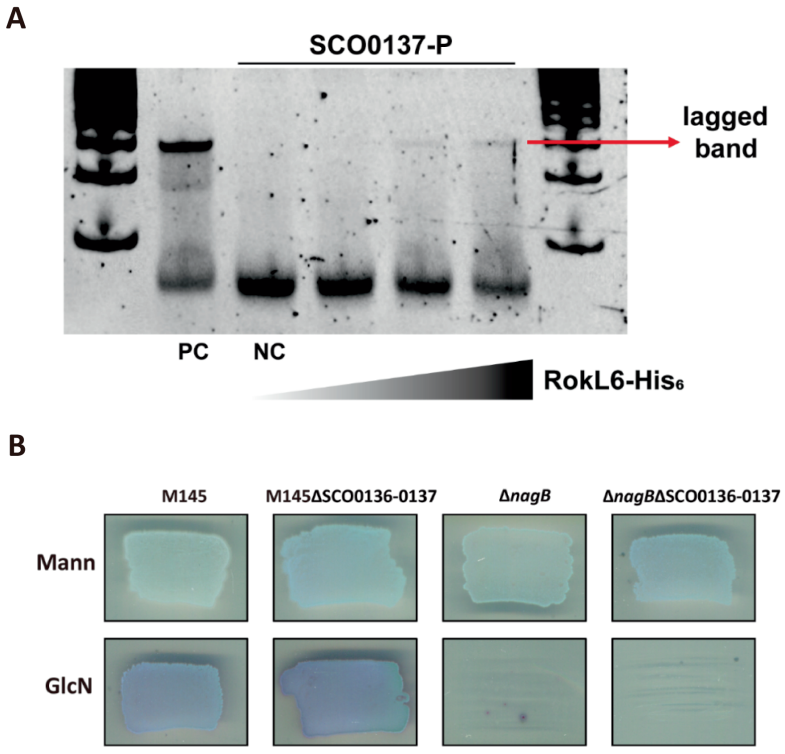

Figure S6. Binding of RokL6 to the SCO0137 promoter region SCO0137-P. (**A**) EMSA of RokL6-His_6 with the upstream region of the SCO0137-SCO0136 operon. The *rokL6* promoter region with 0.5 µM RokL6-His_6 was used as positive control (PC), and SCO0137-P without RokL6-His_6 was used as negative control (NC); lagged band is indicated by red arrow. Note that binding signal of RokL6 to the promoter is very weak; (**B**) Growth of SCO0136-0137 mutants on GlcN. *S. coelicolor* M145, ΔSCO0136-0137, Δ*nagB* and Δ*nagB*ΔSCO0136-0137, were grown on MM with mannitol (Mann), and MM with mann and GlcN (GlcN) for their viability evaluation. Note that deletion of SCO0136-0137 in Δ*nagB* did not relieve the toxicity of GlcN.

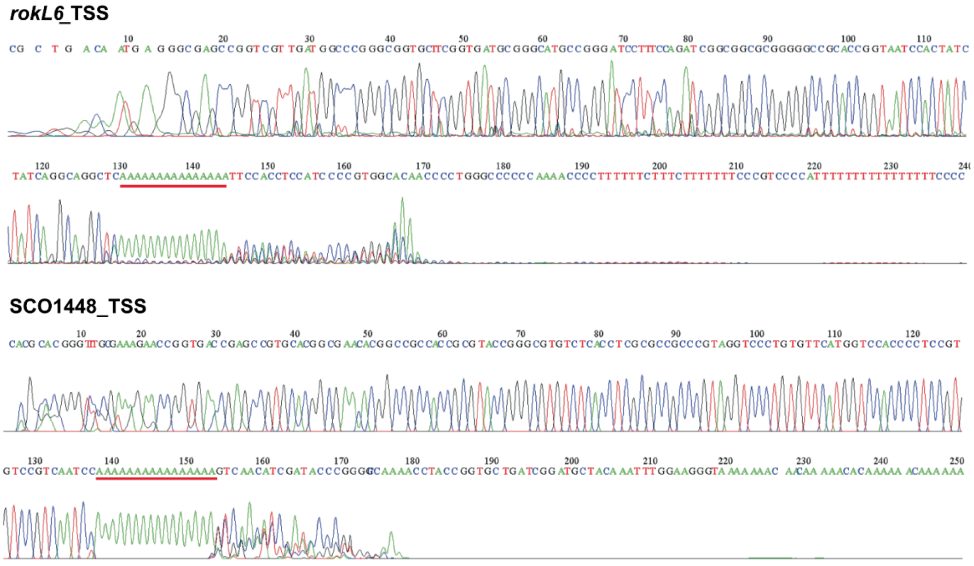

Figure S7. Determination of TSSs of rokL6 and SCO1448 by 5' RACE. Red lines: complementary sequences of oligo(dT) anchor primer.