Global control of development and antibiotic production by nutrient-responsive signalling pathways in Streptomyces

Swiatek, M.A.

Chapter V

System-wide *in vivo* analysis of the regulon for the nutrient sensor DasR in *Streptomyces coelicolor* reveals its global role in controlling metabolism and development

Magdalena A. Świątek, Giselda Bucca, Emma Laing, Fritz Titgemeyer, Colin P. Smith, Sébastien Rigali and Gilles P. van Wezel

submitted for publication

ABSTRACT

Streptomycetes produce a wealth of natural products, including antibiotics, anti-cancer agents and immunosuppressants. *Streptomyces* genomes likely harbour many sleeping antibiotic gene clusters, and to allow exploitation of these hidden treasures, we need to understand the global mechanisms that control their expression. Here we show by an integrated systems level analysis that the nutrient sensor DasR is a master regulator of primary and secondary metabolism and development, with extensive primary and secondary response regulons. DasR binding *in vivo* is relieved by extracellular addition of *N*-acetylglucosamine. Two classes of direct targets were identified, namely those associated with a DasR-responsive element (*dre*) that were also bound *in vitro* (Class I) and those not associated with a *dre*, likely requiring a co-repressor for binding (Class II). In line with the function of DasR as a master regulator, additional specific induction is typically required for the induction of specific regulons. The control of early developmental regulatory genes such as *bldD*, *bldK*, *bldN*, *crp* and *wblA* suggests that DasR is one of the earliest acting regulatory networks in the *Streptomyces* life-cycle.

INTRODUCTION

The ability of bacteria to adapt to a wide range of nutritional environments requires intricate regulatory systems to ensure that gene expression connects to changes in the key routes of primary metabolism. The flux of carbon and nitrogen is thereby typically controlled by a number of global regulators, each with a larger number of locally acting regulators in their response regulon. The global regulators ensure that the flux of essential metabolites is closely monitored and tuned (Martinez-Antonio & Collado-Vides, 2003; Sonenshein, 2007). Streptomycetes are soil-bound bacteria and adaptation to the rapidly changing conditions in this habitat is particularly challenging and hence likely requires a network of sensory systems. Indeed, the *Streptomyces coelicolor* genome encodes 45 ECF (extracellular functions) sigma factors and some 70 two-component regulatory systems (Hutchings *et al.*, 2004). As producers of over half of the known antibiotics, streptomycetes are a paradigm of secondary metabolite-producing microorganisms (Chater & Losick, 1997; Hopwood, 2007). Typically, antibiotic biosynthetic gene clusters are activated by a pathway-specific regulator within or adjacent to the cluster, and in a growth-phase dependent manner (van Wezel & McDowall, 2011). A pivotal question to address is how the primary environmental signals are relayed to global carbon and nitrogen control and further to the control of development and antibiotic production. The elucidation of the *S. coelicolor* genome sequence - the first from an antibiotic-producing actinomycete (Bentley *et al.*, 2002) – was a landmark event, and showed that the antibiotic-producing potential of actinomycetes had been underestimated. Indeed, actinomycete genomes typically contain some 20 sets of putative biosynthetic genes for secondary metabolites, and it is yet unclear how many specific biologically active antibiotics (Challis & Hopwood, 2003). Many of these clusters are not expressed under typical laboratory conditions of rapid growth on nutrient-rich media, often referred to as sleeping or cryptic antibiotic clusters. This may explain why they have been missed in the massive screening efforts performed by Big Pharma (Cooper & Shlaes, 2011; Payne *et al.*, 2007). Many antibiotic biosynthetic gene clusters are regulated in subtle ways so as to allow the organisms to adapt to life in the soil, with its varying physical, chemical and biological stresses, and the challenge is now to exploit this untapped source of potentially valuable natural products (Baltz, 2008; Hopwood, 2007).

To allow activation and screening of sleeping antibiotic biosynthetic clusters,

detailed insight is required into the linkages between environmental (nutritional) signals and secondary metabolite production. Carbon source utilization is a major determining factor in the metabolic control of antibiotic production (Sanchez *et al.*, 2010; van Wezel & McDowall, 2011). The major control system for carbon utilization in bacteria is carbon catabolite repression (CCR). In the model organisms *Escherichia coli* and *Bacillus subtilis*, CCR is mediated via the phosphoenolpyruvate-dependent phosphotransferase system (PTS; reviewed in Brückner & Titgemeyer, 2002). In streptomycetes, the PTS does not seem to play a major role in carbon control and CCR is mediated through glucose kinase, although the mechanism is still unknown (Angell *et al.*, 1994; van Wezel *et al.*, 2007). Additionally, glucose is not internalised via the PTS but rather via the major facilitator GlcP permease (van Wezel *et al.*, 2005). In streptomycetes the PTS fulfills a dual role, and is involved in both *N*-acetylglucosamine (GlcNAc) uptake (Nothaft *et al.*, 2003; Nothaft *et al.*, 2010) as well as in the switch from normal growth to development (Rigali *et al.*, 2006; Rigali *et al.*, 2008). This switch is controlled by the extracellular level of GlcNAc, accumulation of which causes developmental arrest and blockage of antibiotic production in *S. coelicolor* (Rigali *et al.*, 2006). GlcNAc is the monomer of the abundant natural polymer chitin - after cellulose the second most abundant carbohydrate on earth - and also a major constituent of cell-wall peptidoglycan. GlcNAc is a privileged C- and N-source for streptomycetes, and the related metabolite glutamate is even preferred as a carbon source over glucose by *S. coelicolor* (van Wezel *et al.*, 2006). The GlcNAc regulon is controlled by the GntR-regulator DasR (Rigali *et al.*, 2006; Rigali *et al.*, 2004; Świątek *et al.*, 2012), and DNA binding activity is inhibited *in vitro* by glucosamine-6-phosphate (Rigali *et al.*, 2006), a metabolic product derived from GlcNAc. DasR was identified as a nutrient sensory protein that is not only essential for development of streptomycetes, but also directly controls antibiotic production. *In silico* and *in vitro* evidence revealed likely direct repression of the pathway-specific activator genes for *actII*-ORF4 and *redZ* for actinorhodin and prodiginine production, respectively (Rigali *et al.*, 2008). The effect of GlcNAc on *S. coelicolor* depends on the specific culture conditions. On rich media, addition of GlcNAc blocks development and antibiotic production, whereas on minimal media with agar as the sole carbon source it accelerates these processes. These different media are considered to reflect the states of feast or famine, respectively, in the natural environment.

Here we describe a system-wide analysis of the DasR regulon and show that it plays a key role in the control of genes necessary for primary metabolism and antibiotic production, as well as for control of early growth and development. The study also reveals how DasR binding varies with culturing conditions, with relieve of *in vivo* binding following the induction by GlcNAc.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Escherichia coli JM109 (Sambrook, 1989) and ET12567 (Kieser *et al.*, 2000) were used for routine cloning procedures and for extracting non-methylated DNA, respectively. All media and routine *Streptomyces* techniques are described in the *Streptomyces* manual (Kieser *et al.*, 2000). Cells of *E. coli* were grown in Luria–Bertani broth (LB) at 37°C. SFM (mannitol soya flour) agar plates were used to prepare spore suspensions.

Strains were cultivated on MM agar plates with mannitol (1% w/v) (covered with cellophane discs) or in liquid R5 or NMMP supplemented with mannitol (1% w/v). For induction experiments, liquid cultures were grown at 30°C until mid-logarithmic phase (OD_{600} approximately 0.4 for NMMP cultures and 0.8 for R5 cultures), followed by the addition of GlcNAc (50 mM). For the ChIP-on-chip studies on MM mannitol solid medium mycelium was harvested after 24 h and 54 h.

Plasmids and constructs

For the ChIP-on-chip studies a recombinant *dasR* gene was synthesised that allows production of DasR containing a C-terminal triple FLAG-tag epitope (GenScript, La Jolla, USA). The DNA fragment contained the *dasR* gene along with 306 bp promoter region, with the stop codon of the gene replaced by the sequence 5'-ATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTAC AAGGACGACGACGACAAGTAGACCAGAGCCCGCTCACCCGGCCCCAGATTG CGGTTGAAGTCC-3' (stop codon underlined). In this way, the DasR protein was extended with the amino acid sequence MDYKDHDGDYKDHDIDYKDDDDK. The fragment was cloned into the integrative vector pSET152 (Bierman *et al*, 1992). This resulted in vector pGAM29, which expresses DasR-3xFLAG for use in ChIP-on-chip experiments.

Chromatin immunoprecipitation, gDNA labeling and array hybridizations

S. coelicolor M145 containing the integrative vector pGAM29 expressing 3xFLAG-tagged DasR and the *dasR* null mutant GAM29 were grown in duplicate (see *Bacterial strains and growth conditions*). Cultures were treated with formaldehyde (final

concentration 1%, 10 min, 30ºC) in order to crosslink proteins to DNA. Subsequently glycine (final concentration of 0.5M) was added to quench the formaldehyde and cultures were incubated for further 5 minutes at 30ºC. Mycelium was harvested by centrifugation, washed twice with 1xPBS and stored at -70ºC. Chromatin isolation, labeling and hybridization were performed as described previously (Bucca *et al.*, 2003; Bucca *et al.*, 2009). Biological replicates of DasR-3FLAG-IP chromatin were labeled with either Cy3 or Cy5 –dCTP in a dye swap experimental design and co-hybridised with the mock 'no-antibody' IP chromatin on 4 ´ 44K whole genome *Streptomyces* arrays manufactured by Agilent Technologies (Bucca *et al*., 2009).

RNA isolation, cDNA synthesis and labeling

Total RNA was purified using the Kirby-mix protocol (Kieser *et al*., 2000). DNaseI treatment was used to fully remove any traces of DNA. Before use, the RNA preparations were checked for their quality and integrity on the Agilent 2100 Bioanalyzer (Agilent Technologies). cDNA synthesis and labeling was performed as previously described (Bucca *et al.*, 2003; Bucca *et al*., 2009).

Microarray expression data and ChIP-on-chip data analysis

For the time-course gene expression studies *S. coelicolor* M145 and *dasR* mutant were cultivated on MM mannitol agar plates overlaid with cellophane discs. Samples were taken at 24 h, 30 h, 36 h, 42 h and 54 h. Two biological replicates were analysed for all samples and for each strain. The filtered data sets for the gene expression experiments were analysed using Rank Products analysis (Breitling *et al.*, 2004) via the Rank-Prod package in R (version 2.5.0) (Hong *et al.*, 2006) using the web-based implementation, RankProdIt Laing E, Smith CP. (2010) 'RankProdIt: A web-interactive Rank Products analysis tool.' *BMC Res Notes*, 3 221). Differentially expressed genes were identified as having a pfp (probability of false prediction) value (Breitling *et al.*, 2004) less than or equal to 0.15, equal to a false discovery rate of approximately 15%. Some additional genes were also identified that demonstrated a >1.5 fold change in expression between wild-type and the *dasR* mutant. ChIP-on-chip data, allowing identification of genes/ regions directly controlled by DasR were analysed as described previously (Bucca *et al.*, 2009).

RESULTS

Predicted DasR regulon and growth conditions

Previous bioinformatic predictions using the PREDetector algorithm (Hiard *et al.*, 2007) revealed some 200 sequences that conformed to the consensus binding site for DasR (*dre*, for DasR responsive element), namely the palindromic 16 bp consensus sequence A (G/C)TGGTCTAGACCA(G/C)T. These sequences are found primarily associated with genes related to primary metabolism, sugar transport and extracellular polysaccharide hydrolysis, as well as antibiotic production (Craig *et al.*, 2012). The transcriptomic analysis of the mutant of the pleiotropic *dasR* regulatory gene, with impact on primary metabolism, secondary metabolite production and sporulation, requires a well-defined experimental set up. Major changes were previously observed in the *dasR* mutant in terms of antibiotic production on either MM agar plates with only agar as the carbon source, or on complex media (Rigali *et al.*, 2008). However, growth is poor on MM with agar and *dasR* null mutants fail to sporulate on rich media, and strains in different phases of the life cycle can not be properly compared (den Hengst *et al.*, 2010). We therefore used MM agar plates with mannitol for comparison of wild-type and *dasR* mutant, as strains display roughly similar development under these growth conditions (Rigali *et al.*, 2006). RNA was isolated from mycelia harvested after 24 h, 30 h, 36 h, 42 h or 54 h of growth, corresponding to vegetative growth, early aerial growth, aerial growth, early sporulation and sporulation, respectively.

Transcriptomic analysis

Microarray analysis revealed some 1200 genes that were significantly differentially expressed in the *dasR* null mutant of *S. coelicolor* relative to the parent M145, and the transcription of over 400 genes was more than two-fold up- or down-regulated at one or more time points in the *dasR* mutant (across two biological replicates). These genes were considered as significantly affected by the deletion of *dasR* (Table S1); of these, some 100 genes showed more than three-fold and 26 genes showed more than five-fold change in transcription relative to the parental strain. Classes of genes that were overrepresented are those related to: primary metabolism; protein biosynthesis; development and secondary metabolism; SapB, chaplins, rodlins and other small peptides; transposases

and DNA recombination.

Primary metabolism and protein biosynthesis

Genes of the DasR core regulon (*nag, pts*) were up-regulated in the *dasR* mutant and mostly at all time points, including *nagE2*, *nagF*, *ptsH (*encoding PTS components*); nagB*, *nagK* (encoding GlcNAc metabolic enzymes)*; glmS1* and *glmS2* for glucosamine-fructose-6-phosphate aminotransferase (which has the opposite catalytic activity as NagB); and *dasA-dasBC* (for the chitobiose transporter (Colson *et al.*, 2008; Saito *et al.*, 2007) (Fig. 1A). Most of these genes are direct targets, and hence repressed by DasR (Colson *et al*., 2007; Rigali *et al*., 2008; Świątek *et al*., 2012). Additionally, several genes involved in central metabolism, such as glycolysis, the TCA cycle, gluconeogenesis and pyruvate metabolism as well as for biosynthetic pathways for amino acids, nucleotides and fatty acids were up-regulated in the *dasR* mutant (Supplemental Table S1). Strong targets were ABC transporters, in particular the branched-chain amino acid transporter operon SCO2008-SCO2012, the SCO6005-SCO6007 operon (Fig. 1A) that is orthologous to the *ngcEFG* chitobiose/GlcNAc transport operon in *S. olivaceoviridis* (Wang *et al.*, 2002) and the oligopeptide transporter operon SCO5476-SCO5480 (Supplemental Table S1). Despite the strong binding of DasR to the genes belonging to the chitinolytic system (see below), their transcription was similarly low in both wild-type and *dasR* mutant cells. The lack of enhanced expression of the *chi* genes is most likely explained by the lack of inducer that is required to activate their transcription.

The widespread deregulation of protein biosynthetic genes is noteworthy, in particular almost all genes for ribosomal proteins, as well as genes for translation initiation factors IF-1 and IF-3 and elongation factors EF-G, EF-P, EF-Ts, EF-Tu1 and EF-Tu3, as well as the chaperonins GroES, GroEL1 and GroEL2. Transcription of all of the genes belonging to the larger ribosomal protein operons S10 (SCO4701-4711), L14 (*spc*; SCO4712-4722), S12 (*str*; SCO4659-4662) and S13 (SCO4727-4730) responded in the same way to the deletion of *dasR*, with up-regulation during early growth (24 h) and sporulation (54 h), while the genes were down-regulated around 30 h (change from vegetative to aerial growth). No major changes were found at 36 and 42 h (aerial growth) (Supplemental Fig. S1A).

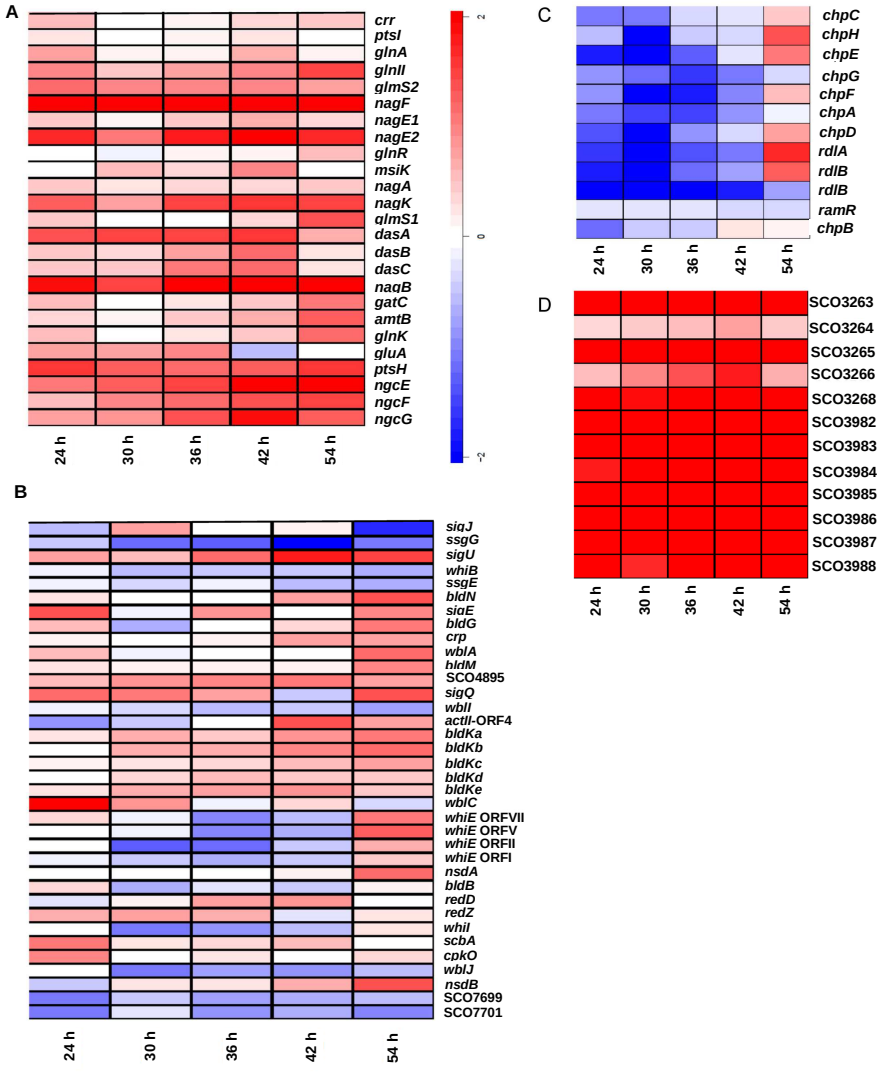


Figure 1. Selected categories of genes differentially expressed in the *dasR* null mutant. Important categories of genes whose transcription was altered more than two-fold in microarray experiments comparing the *dasR* mutant GAM29 relative to its parent *S. coelicolor* M145 are shown; these are (A) genes for amino acid and (amino-)sugar transport and metabolism; (B) genes related to development and antibiotic production, (C) genes for chaplins, rodlins and SapB, and (D) the SCO3263-3268 and SCO3982-3988 operons for small peptides of unknown function. RNA was isolated from *S. coelicolor* M145 or its *dasR* mutant GAM29 cultivated on MM mannitol agar plates after 24 h (vegetative growth), 30 h (vegetative/aerial growth), 36 h (aerial growth), 42 h (aerial growth/sporulation) and 54 h (sporulation). A heat map is provided as reference for the fold changes in log scale ($\log_2$(mutant/wild-type)). Note that many genes shown a larger than four-fold change in transcription. For exact values refer to Supplemental Table S1.

Secondary metabolism

Transcription of the biosynthetic gene clusters for actinorhodin (SCO5070-SCO5091) and for prodiginine was enhanced in the *dasR* mutant, in particular around 42 h (Supplemental Table S1). Interestingly, *scbA* (SCO6266), which encodes the γ-butyrolactone synthase ScbA implicated in the control of the antibiotic clusters *act, red* and *cpk* (encoding a cryptic modular type I polyketides synthases) in *S. coelicolor* (Hsiao *et al*., 2007), was up-regulated at 24 h (Fig. 1B). *cpkO* (SCO6280), the transcriptional activator gene of the *cpk* cluster, was also up-regulated at 24 h, and this is followed by enhanced expression of the *cpk* cluster in the next time point (30 h). Conversely, four genes (SCO3218, SCO3222, SCO3224, SCO3244) belonging to the *cda* cluster for calcium-dependent antibiotic CDA were down-regulated, in particular at 24 h (Fig. 1B and Supplemental Table S1).

Transcription of the two-component system genes *afsQ1-Q2* (SCO4906-07) was also enhanced. In *S. coelicolor*, AfsQ activates antibiotic production on MM with glutamate as sole nitrogen source (Shu *et al.*, 2009). Other antibiotic-related genes that were up-regulated in the *dasR* mutant include *rfsA* (SCO4677), an *abaA*-like regulatory gene which controls *afsQ* and represses antibiotic production and morphological differentiation (Choi *et al*., 2010; Kim *et al*., 2008), *nsdA* (SCO5582), a negative regulator of antibiotic production and morphological differentiation in *S. coelicolor* (*Li et al*., 2006) and its paralogue *nsdB* (SCO7252; Zhang *et al*., 2007).

Finally, the deoxysugar gene cluster SCO0381-SCO0401 was up-regulated in the *dasR* mutant (Supplemental Table S1). SCO0381-SCO0401 shows significant similarity to the *ste* gene cluster (*Streptomyces* eps) of *Streptomyces* sp. 139 required for Ebosin biosynthesis. Ebosin is an exopolysaccharide (EPS) which shows anti-rheumatic arthritis activity *in vivo* (Wang *et al*., 2003).

Development

The DasR regulon includes a surprisingly large number of regulators, with 58 regulatory genes significantly up- or down-regulated in the mutant at one or more time points. In line with the notion that DasR acts in particular as a switch from vegetative to aerial growth, specifically early developmental genes were up-regulated in the *dasR* mutant at later time points, namely *bldK, bldM, bldN, crp* and *wblA* (Fig. 1B), while expression of

sporulation genes (*e.g. whi, ssg*) was left unchanged. The pentacistronic *bldK* operon encodes an oligopeptide transporter required for early stages of development (Nodwell & Losick, 1998; Nodwell *et al.*, 1996), *bldN* encodes a developmentally regulated σ factor, σ^{BldN}, required for aerial mycelium formation, which in turn is required for the transcription of the response regulator gene *bldM* (Bibb *et al*., 2000), *crp* is involved in the control of spore germination (Derouaux *et al*., 2004; Piette *et al*., 2005), and the WhiB-like protein WblA represses antibiotic production and may also be involved in the transition from early aerial hyphal cells to the subapical stem and apical compartments that precede sporulation (Fowler-Goldsworthy *et al*., 2011). Some of the most highly up-regulated genes were *wblC* (SCO5190) and its flanking genes SCO5189 and SCO5191, and *mtrA* (SCO3013), the homologue of which controls proliferation of *Mycobacterium tuberculosis* (Fol *et al*., 2006). Expression of *ssgG* (SCO2924), a homologue of *ssgB* whose gene product recruits FtsZ during sporulation (Willemse *et al.*, 2011), was down-regulated in the *dasR* mutant (Fig. 1B).

Strikingly, *chpABCDEFGH, rdlAB,* and *ramS*, encoding chaplins, rodlins and SapB, respectively, were massively down-regulated during early time points. The chaplins and rodlins are hydrophobic 'coat' proteins that assemble into a rodlet layer that provides surface hydrophobicity to aerial hyphae and spores, thus allowing them to break through the moist soil surface (Claessen *et al*., 2003; Claessen *et al.*, 2006; Wösten & Willey, 2000); the spore-associated protein SapB is a lantibiotic-type signalling peptide required for the onset of development (Willey *et al*., 1991). Several of the genes featured in the top 10 of most strongly down-regulated genes, particularly during early time points (Fig. 1C), strongly suggesting their direct or indirect dependence on DasR. The absence of *dre* elements associated with the genes, as well as lack of direct binding in ChIP-on-chip experiments (see below) strongly suggests that a transcriptional regulator for these genes is deregulated in the *dasR* null mutant.

Transcription of two other gene clusters encoding small (60-100 aa) secreted proteins and a GntR regulator, was also strongly repressed, with transcription of the SCO3982-SCO3988 cluster so strongly enhanced that it suggests that the cluster is silenced by DasR (Fig. 1D). Transcription of the paralogous gene cluster SCO3263-3268 (with SCO3982 similar to SCO3268 and so on) was also strongly repressed by DasR (Fig. 1D). SCO3983 contains a TTA codon suggesting it is subjected to translational

control by the tRNA BldA, which effects developmental control of gene expression (Lawlor *et al*., 1987). The function of the gene clusters is yet unknown, but their massive up-regulation in the *dasR* mutant is intriguing.

Identification of the direct DasR binding sites *in vivo* by ChIP-on-chip

In order to discriminate between genes directly or indirectly regulated by DasR and to obtain detailed biological insight into the genome-wide distribution of DasR binding sites *in vivo*, we conducted ChIP-on-chip analysis using mycelia grown on MM agar plates with mannitol as the sole carbon source. For this, the binding profiles of wild-type *S. coelicolor* M145 carrying the integrative vector pGAM29, which expresses C-terminally 3xFLAG-tagged DasR, and the *dasR* null mutant GAM29 (control) were directly compared. For each strain two biological replicates were analysed. Samples were collected at 24 h and 54 h, which for the parental strain corresponded to vegetative growth and full development (aerial hyphae and spores), respectively. The *dasR* null mutant also sporulated under these conditions, and growth rates were essentially the same for the two strains. DasR showed statistically relevant binding to 76 sites after 24 h and to 265 sites after 54 h (Supplemental Table S2). However, using a stringent threshold of 0.5 (log2 (wild-type with *dasR*-flag antibody/no-antibody)/(*dasR* mutant antibody/no-antibody)) to identify significantly enriched probes (following a t-test as detailed in Bucca *et al*., 2009), and setting a minimum of two adjacent probes significantly enriched, this list was reduced to 22 and 57 sites that had their promoter proximal regions bound with high probability in the 24 h and 54 h old samples, respectively (Table 1). The distribution of the *in-silico* predicted and experimentally validated DasR-binding *dre* elements is presented schematically in Fig. 4A and Fig. 4B, respectively. It should be noted at this point that we also performed eight additional ChIP-on-chip experiments on liquid-grown MM and R5 cultures, to analyse the effect of GlcNAc on *in vivo* binding (see below). Taken together, this provided very good validation for the *in vivo* binding of DasR, with up to eight independent verifications for the target genes (Table 1).

Table 1. Summary of ChIP-on-chip experiments and validation by EMSAs (+, binding; -, no binding)

Gene ID or nearest gene	Name	24h	54h	Induction		IDs #	EMSA	*dre* (cis-)element		
				MM	R5			Sequence 5'-> 3'	position "	score ^
Primary metabolism										
SCO1390*-SCO1391*	*crr-ptsI*	+	+	+	+	8	+	tgtggtctagacctct	-130	17.14
SCO1429*[54h]	*chiD*	+	+	-	+	4	+	actggtctagtcctcc	-96	12.17
								aatggtccgaaccatt	-118	7.45
SCO1444*	*chiI*	+	+	-	-	2	+	actggtctagtcctct	-53	16.23
								attggtccatacctat	-75	10.22
SCO2503*[54h]	*chiJ*	+	+	-	+	4		aaaggtctggaccaca	-78	12.37
								cttggtccagacctct	-99	10.78
								tctggaccacagcact	-73	7.29
SCO2833*	*chb*	+	+	+	+	6		acatgtccataccaaa	-110	9.34
								gcaggtctagaccaag	-70	8.67
SCO2906	*nagF*	+	+	-	+	6		actggtctacaccagt	-41	17.52
								atcggtctgcaccagt	718	8.37
								caaggtgtagacctct	1287	12.27
SCO2907	*nagE2*	+	+	-	+	4		acaggtctacaccact	-49	16.57
								agtggtgtagaccacc	-32	14.37
								caaggtgtagacctct	-236	12.27
SCO4284-SCO4285	*nagKA*	+	+	+	+	7	+	agaggtctagtccact	-83	16.64
								ggtggtgtagacctta	-63	9.64
SCO5003-SCO5004*[54h]	*chiA*	+	+	+	+	4		ggtggtccagaccaat	-77	10.43
								ggtggtccagaccaat	-258	10.43
SCO5230		+	+	+	+	5		tctggtctagtcctgg	-118	9.87
SCO5231-5232*	*dasR*	+	+	+	+	5	+	cttggtctagtccata	-150	9.05
								tctggtctagtcctgg	755	9.87
	dasA						+	actggtctacaccatt	-106	17.13
								cttggtctagtccata	-322	9.05
SCO5236	*nagB*	+	-	-	+	3	+	tgtggtttagaccaat	-68	17.51
SCO5376	*chiC*	+	-	-	-	1		ataggtctggaccaat	-109	11.92
								aaaggtctggaccata	-88	9.23
SCO5423	*pyk2*	-	-	+	-	4				
SCO5673*	*chiB*	+	+	+	+	7		attggtctggaccaaa	-63	10.69
SCO5841	*ptsH*	+	+	-	-	2	+	agttgtctagaccagt	-51	18.11
								tcttgtctagaccagt	-66	16.25
SCO6004-SCO6005*[54h]	*ngcE*	-	+	+	+	7		agtggactatacctgt	-244	16.05
								agtggactatacctgt	-334	16.05
SCO6012-SCO6013	*chiH*	+	-	-	+	2	+	aatggtctggaccaga	-111	12
								atgggactagaccaat	-127	10.22
								aatggtctggaccaga	-274	12
								atgggactagaccaat	-258	10.22
SCO6032-33		+	+	-	+	4		cttggtctagtccatt	-154	10
								cttggtctagtccatt	-278	10
SCO6300		+	-	-	+	2		ataggtctagacaaaa	-131	13.72
								agaggtctagacaaaa	-116	13.48
SCO6344-45		+	-	-	+	2		taaggtctagacctgc	-133	9.6
								gtaggtctagacctgc	-153	8.13
	chi	+	-	-	+	2	+	taaggtctagacctgc	-114	9.6
								gtaggtctagacctgc	-94	8.13
SCO7225*[54h]	*chi*	+	+	-	+	3	+	tcaggtctagacctgt	-34	15.28
								ccttgtctagaccaat	-168	13.38
								tatggtctagacctga	-55	13.25
SCO7335[c]	*pep1B*	-	+	-	-	1				

Development and secondary metabolism										
SCO3230[c]	*cdaPSI*	-	+	+	+	5				
SCO3231[c,*54h]	*cdaPSII*	-	+	+	+	5				
SCO5085	*actII-4*	-	-	+	+	4	+	tgttgagtaggcctgt	-59	10.49
SCO5877	*redD*	-	-	+	+	4		actgctggagaccggt	612	7.15
SCO5881[c]	*redZ*	-	-	+	+	4	+	agtggtttccacctca	-201	12.44
SCO6273[c]	*cpkC*	-	+	+	-	4		acatgcgtaatcaact	-13	9.2
SCO6274[c]	*cpkB*	-	+	+	-	4				
SCO6275[c,*54h]	*cpkA*	-	+	-	-	1				
SCO1276	*sigJ*	-	+							
SCO1488-89[*54h]	*pyrR*	-	+	-	+	5				
	bldD									
SCO3323[*54h]	*bldN*	-	+	-	-	1				
Translation										
SCO2504-05	*glyS*	-	+	-	-	1		agtggtctgcacctgg	515	10.75
SCO3092		+	+	-	+	5				
SCO3679		+	+	-	-	2		tgttgtctagtccaat	-314	15.32
SCO4092		-	-	-	+	4				
SCO5239		+	+	-	-	2	+	agtggtctagtccaca	-335	16.23
SCO5550[d]		+	+	+	+	8	+	actggtctaaaccaca	-18	17.27
SCO4123[d]		-	+	+	+	6				
SCO1792[d]		-	+	+	+	6				
SCO1390		-	+	+	+	6				
SCO3019[d]		-	+	+	+	6				
SCO5746		-	+	+	+	6				
SCO3334		-	+	+	+	6				
Transposases and DNA topology										
SCO0020[c]		-	+	-	-	1				
SCO0098-SCO0099		-	-	+	-	2				
SCO0567		-	-	+	-	3				
SCO1471		-	-	+	-	1				
SCO1603		-	-	+	-	3				
SCO3466		-	-	+	+	5				
SCO3467		-	-	+	-	3				
SCO3468		-	+	-	-	1				
SCO3490[c]		-	+	-	-	1		aatcgtcaagacctgt	-118	8.75
SCO4344		-	+	-	-	1				
SCO4698[c]		-	+	-	-	1				
SCO4699[c]		-	+	+	+	4		agaggtcgacacccgt	2130	9.17
SCO6400[c]		-	+	+	+	6				
SCO6627[c]	*pglX*	-	+	-	-	1		cacggtgtagacatca	841	7.41
SCO6910-SCO6911		-	-	+	+	5				
SCO7798[c]		-	+	-	-	1				
Other										
SCO1606		-	-	+	+	5				
SCO2078		-	+	+	+	6				
SCO3262[c]		-	+	+	+	4		acaggaggacaccatt	-118	8.53
SCO5190	*wblC*	-	-	+	-	4				
SCO7056		+	+	-	+	3		attggtctaaaccagc	-79	15.65
SCO4646	*secE*	-	-	-	+	4				
SCO4067	*dnaZ*	-	-	-	+	3				

* one probe; * 54 h one probe at 54 h; c coding; d downstream

Total nr of independent ChIP-chip experiments showing DasR binding

^ matrix score indicating the statistical relevance of the predicted *dre* element relative to the consensus sequence (only scores >7 were considered)

" position indicates spacing between the dre element and start of the gene

The nag/chi regulon and validation of the ChIP-on-chip data

Strong signals for DasR binding were found in the upstream regions of many genes that belong to the *nag* regulon, the *pts* genes and the chitinolytic system, which can be described as the core regulon for DasR, and coincide with bioinformatically predicted *dre* elements with generally very high scores (Table 1). These include the upstream regions for the *nag* metabolic genes *nagKA* and *nagB*, the *pts* genes *nagF, nagE2, ptsH, crr-ptsI*, and members of the chitinolytic system, namely *chiA, chiB, chiC, chiD, chiH, chiI, chiJ*, SCO6032 (for *β-N*-acetylhexosaminidase), SCO6300 (for a secreted *β-N*-acetyl -glucosaminidase), SCO2833, SCO6345, the SCO6005-6007 operon, *dasA* and *dasR* itself (Fig. 2A and Supplemental Fig. S2). It should be noted that *crr-ptsI*, *chiI*, *chiB*, SCO2833, the SCO6005-6007 operon promoter and *dasR* only had one probe significantly enriched, but most of these were also bound in the induction experiments (Table 1). Validation for the biological relevance of these ChIP-on-chip data is provided by the fact that direct binding of DasR to many genes of the *nag* and *chi* regulons was already demonstrated by electrophoretic shift assays (EMSAs), namely for *nagKA* and *nagB, crr, ptsI, ptsH, nagF* and *nagE2* and for the chitinase genes *chiD, chiH, chiI*, SCO6345 and SCO7225 (Colson *et al.*, 2007; Rigali *et al.*, 2006; Świątek *et al.*, 2012). Importantly, the genomic distribution of the observed DasR binding was different between the cultures harvested at 24 h and 54 h. All of the genes discussed above except SCO6005-6007 were bound by DasR at 24 h, but after 54 h no significant binding was found for *chiC, crr-ptsI, chiH,* SCO6300, SCO6345 and SCO7225. Conversely, binding to the promoter region of the SCO6005-6007 operon was observed at the later point.

Direct control of antibiotic production by DasR

DasR bound exclusively after 54 h to probes corresponding to genes involved in development and antibiotic production. Binding was observed to the intragenic regions of three biosynthetic genes (SCO6273-6275) of the *cpk* gene cluster for the cryptic type I polyketide synthase (Cpk) and of *cdaPSI and cdaPSII* (SCO3230 and SCO3231) encoding the peptide synthetases I and II for the synthesis of calcium-dependent antibiotic (CDA). Additionally, in liquid-grown cultures binding of DasR was detected in the promoter proximal region of *actII*-ORF4 and *redD, redZ*, encoding the pathway-specific activator genes for Act and Red biosynthetic clusters, respectively (see Table 1,

Fig. 3 and below). *dre* elements were identified *in silico* for *actII*-ORF4, *redZ*, *cpkC*. These data candidate DasR as a pleiotropic regulator of antibiotic production.

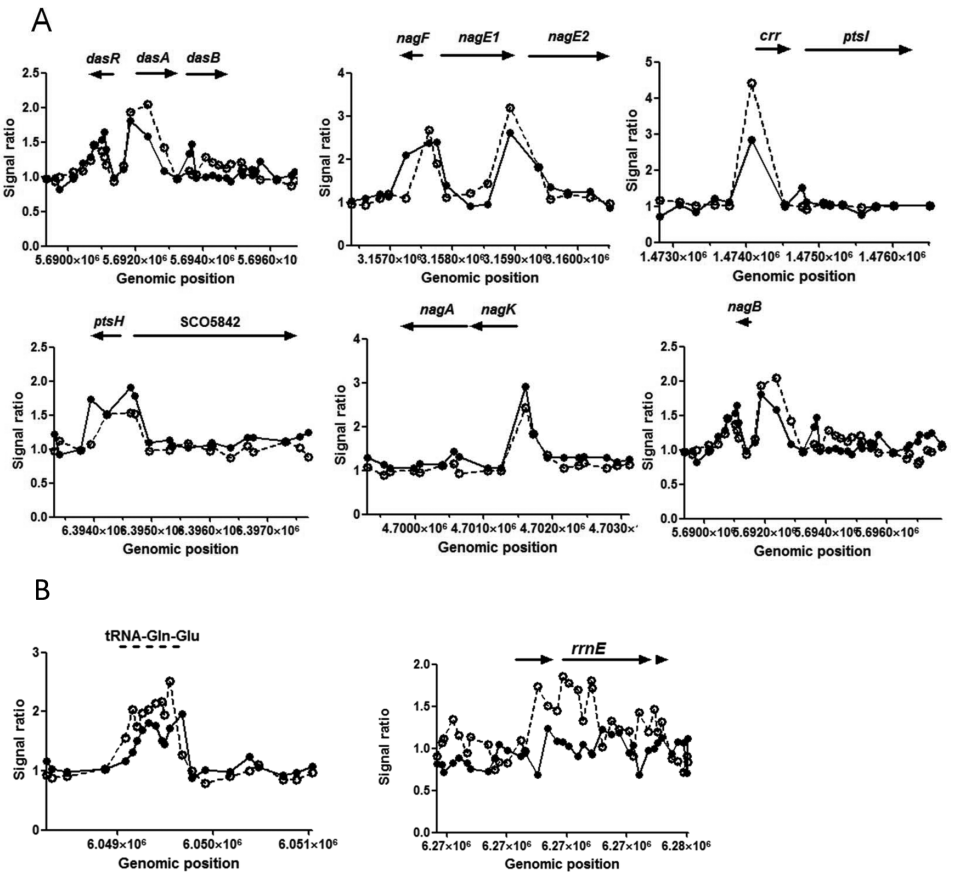


Figure 2. Selected DasR targets identified in ChIP-on-chip time-course experiments. Peaks indicate the presence of DasR binding sites in the regions near (A) the *dasR, dasABC, nag* and *pts* genes (B) the gene cluster specifying $tRNA^{Gln/Glu}$ and the rRNA operon *rrnE*. Plots present DasR binding at 24 h (solid line, closed circles) and 54 h (dashed line, open circles). The arrows indicate the orientation of the target genes. Note that the y-axis (signal ratio) varies between the plots.

Genes for noncoding and stable RNAs

Interestingly, DasR binding was observed for many transfer-RNA (tRNA), ribosomal RNA (rRNA) and noncoding RNA (ncRNA) genes. As for ncRNAs, binding was observed to *dre* elements upstream of scr3092 (Swiercz *et al.*, 2008) and scr5239 (Vockenhuber & Suess, 2012) (Table 1). scr5239 decreases the production of the

antibiotic actinorhodin, and represses expression of the extracellular agarase gene *dagA* (SCO3471) at the post-transcriptional level (Vockenhuber & Suess, 2012). Additionally, DasR likely controls the expression of several tRNA genes, as suggested by binding to multiple sequences associated with highly conserved *dre* elements - and in several independent ChIP-on-chip experiments - within the five-membered $tRNA^{Gln/Glu}$ gene cluster, to a gene for $tRNA^{Gly}$ upstream of SO3679, to $tRNA^{Met}$ and to the $tRNA^{Glu/Asp/Phe}$ gene cluster upstream of SCO4092. Additionally, after 54 h of growth on MM agar plates, DasR binding was observed upstream of SCO2504 (*glyS*) encoding a glycyl-tRNA synthetase (Table 1). Among the most highly enriched probes were those corresponding to the rRNA operons *rrnA, rrnB, rrnC, rrnD, rrnE* and *rrnF* (Table 1 and Fig. 2B), with many probes in the operons bound by DasR. The expression of genes specifying stable RNAs was not detected due to the design of the microarrays (biased for mRNA).

Other targets

Microarray analysis showed massive up-regulation of all genes in the cluster SCO3982-3988, which encode small secreted proteins of unknown function. This gene cluster is most likely directly controlled by DasR, because the GntR regulatory gene SCO3986, the likely regulator of this cluster, was identified as a DasR target in the ChIP-on-chip experiments, both in surface-grown cultures and in liquid-grown cultures (see Table S3 and below).

A surprising number of binding events were detected upstream of or inside transposase genes, namely SCO0020, SCO3468, SCO3469, SCO3490 (and many sites in the genomic region between SCO3468-3490 at 54 h), SCO4344, SCO4698, SCO6400, SCO7798. Several of these genes exist in multiple copies on the *S. coelicolor* genome. Additionally, other DNA recombination related gene such as *pglX* (SCO6627) encoding a putative adenine-specific DNA methyltransferase that is part of a phage integration system was identified as a DasR target. Another genomic region that is overrepresented in terms of the number of binding sites is that between SCO6373-6401 (11 binding sites after 54 h), but most were represented by a single probe and did not result in changes in expression as judged by the microarray data.

Visualization of GlcNAc-mediated induction of the DasR regulon *in vivo*

We previously showed that glucosamine-6P (GlcN-6P) acts as an effector molecule that prevents DasR DNA-binding *in vitro* (Rigali *et al.*, 2006). This and the identification of *dre* elements upstream of all genes for GlcNAc internalization and utilization and the control of antibiotic production, suggested a direct signalling cascade from external nutrient availability to the control of secondary metabolism (Rigali *et al.*, 2006). To visualise and manipulate this putative signalling cascade *in vivo*, we monitored the global response of the DasR regulon to the extracellular addition of GlcNAc. For this purpose, mycelia of *S. coelicolor* M145 harbouring pGAM29 and the *dasR* mutant were each grown in liquid NMMP mannitol or R5 cultures, and cells were induced by the addition of 50 mM GlcNAc. Samples were collected prior to (T_0) and 30 (T_1), 60 (T_2), or 120 min (T_3) after GlcNAc addition, and DasR-enriched chromatin was analysed as described in the Materials and Methods section.

The ChIP-on-chip experiments revealed 54 and 69 sites that were determined as bound by DasR with high statistical relevance in MM- and R5-grown mycelia prior to GlcNAc induction, respectively (Supplemental Table S3). In line with the experiment performed on mycelia from surface-grown cultures, at T_0 DasR was bound to sites belonging to its core regulon *pts*, *nag* and *chi* (Supplemental Fig. S3). Immediately after GlcNAc addition DasR dissociated from the target sites, providing direct evidence that indeed DasR responds to the addition of GlcNAc *in vivo*, thereby reducing its affinity for its target DNA as hypothesised previously (Rigali *et al.*, 2006). Binding was also observed for many transposase-related genes, with dissociation of DasR following addition of GlcNAc. This also connects to possible involvement of DasR in the control of genome rearrangement (see Discussion).

After the addition of GlcNAc, DasR dissociated from the target sites upstream of genes that are part of the deduced GlcNAc-dependent signalling cascade, namely (i) genes for GlcNAc transport and phosphorylation by the PTS (*crr-ptsI* and *nagE2*); (ii) upstream of the transcription unit that includes *nagA* which is necessary to provide the DasR effector GlcN-6P, and (iii) upstream of *actII*-ORF4 and of *redD* and *redZ*, encoding the pathway-specific activator genes for Act and Red biosynthetic clusters, respectively (Fig. 3). Likewise, there was binding in the intragenic region of *cdaPSI* and *cdaPSII* (SCO3230-3231) for CDA peptide synthetases I and II, and for *cpkBC* (for Cpk

polyketide synthases; (Pawlik *et al*., 2007)).

In MM-grown cultures virtually all DasR binding events were relieved following the addition of GlcNAc. In contrast, in nutrient-rich R5 cultures, 29 cases were found where DasR binding was not relieved by the addition of GlcNAc, 15 of which showing enhanced binding. Examples of constant binding include, besides *pyrR-bldD*, the upstream regions of the tRNA genes/operons $tRNA^{Glu/Asp/Phe}$, $tRNA^{Gln/Glu}$ and $tRNA^{Met}$, and *secE* for the preprotein translocase subunit SecE. Examples of enhanced binding of DasR include the upstream regions of *dnaZ* for DNA polymerase II subunit gamma, of *cydA* for cytochrome oxidase subunit I, of the cell division-related gene SCO2078 and of the ncRNA scr3092.

Finally, DasR binding was also observed for *atrA*, and this binding is relieved by the addition of GlcNAc. AtrA is required for actinorhodin production via the direct activation of *actII*-ORF4 (Uguru *et al*., 2005), and also activates the GlcNAc transporter gene *nagE2*, while DasR does the opposite. In the DasR-mediated GlcNAc signalling cascade from nutrient to antibiotic production, internalization via the *nagE2* product is the first step, while control of pathway-specific regulators is the last step. Thus, AtrA and DasR antagonise each other in terms of their influence of this signalling cascade (Nothaft *et al*., 2010).

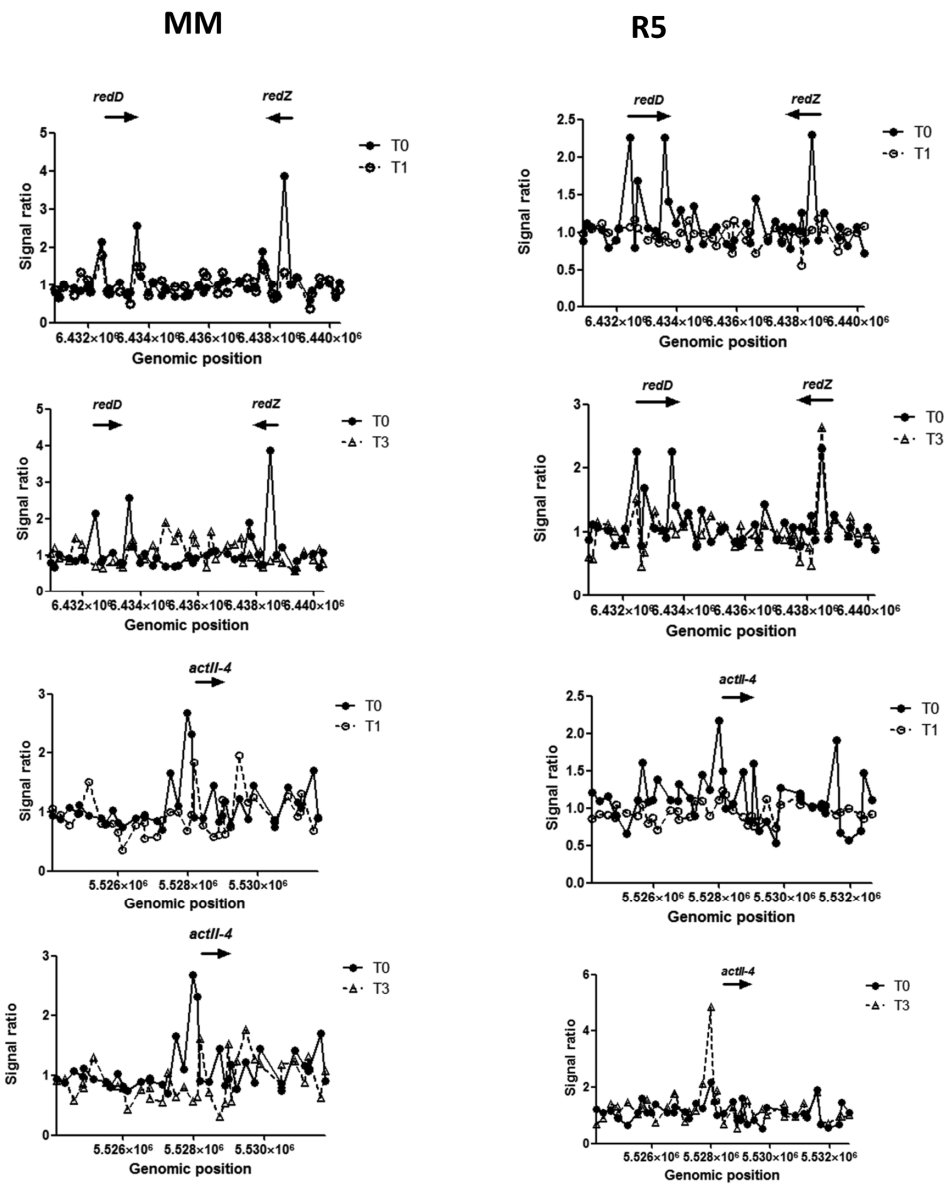
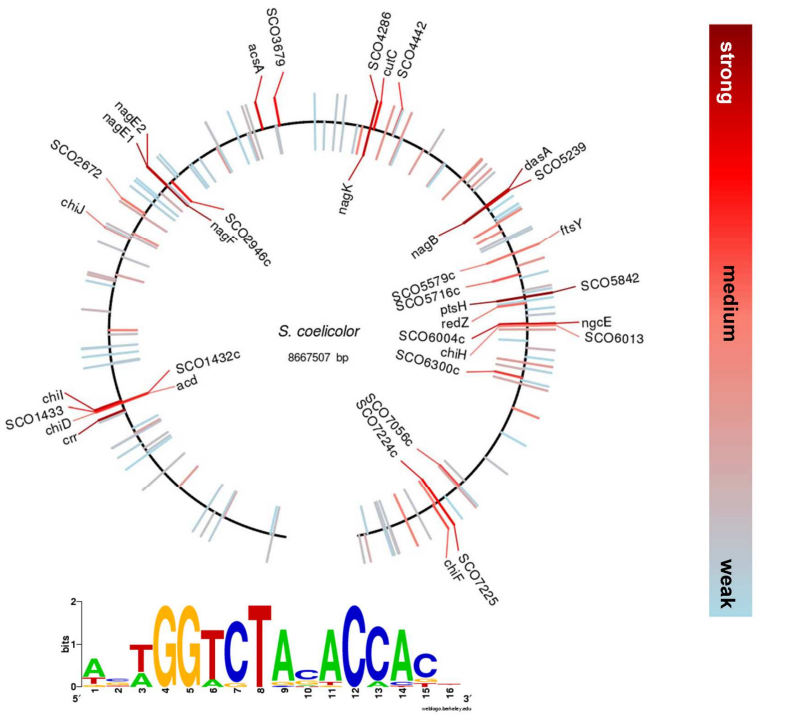


Figure 3. Chip-on-chip data for selected DasR targets detected in the GlcNAc induction experiment. Samples were collected prior to (T_0, closed circles) and 30 (T_1, open circles), 60 (T_2) or 120 min (T_3, open triangle) after GlcNAc addition. Plots indicate change in affinity of DasR for the promoter regions of the pathway-specific activator genes for the biosynthesis of actinorhodin (*actII*-ORF4) and prodiginine (*redD, redZ*) upon addition of GlcNAc (25 mM) to liquid-grown MM (left panel) or R5 (right panel) cultures. Arrows indicate the orientation of the targeted genes. Note that the y-axis (signal ratio) varies between the plots.

A


strong
medium
weak
S. coelicolor
8667507 bp
SCO3679
acsA
SCO4286
cutC
SCO4442
nagK
dasA
SCO5239
nagB
ftsY
SCO5579c
SCO5716c
SCO5842
ptsH
redZ
ngcE
SCO6004c
SCO6013
chiH
SCO6300c
SCO7056c
SCO7224c
chiF
SCO7225
SCO1432c
acd
chiI
SCO1433
chiD
crr
SCO2672
chiJ
nagE2
nagE1
nagF
SCO2946c
bits

B

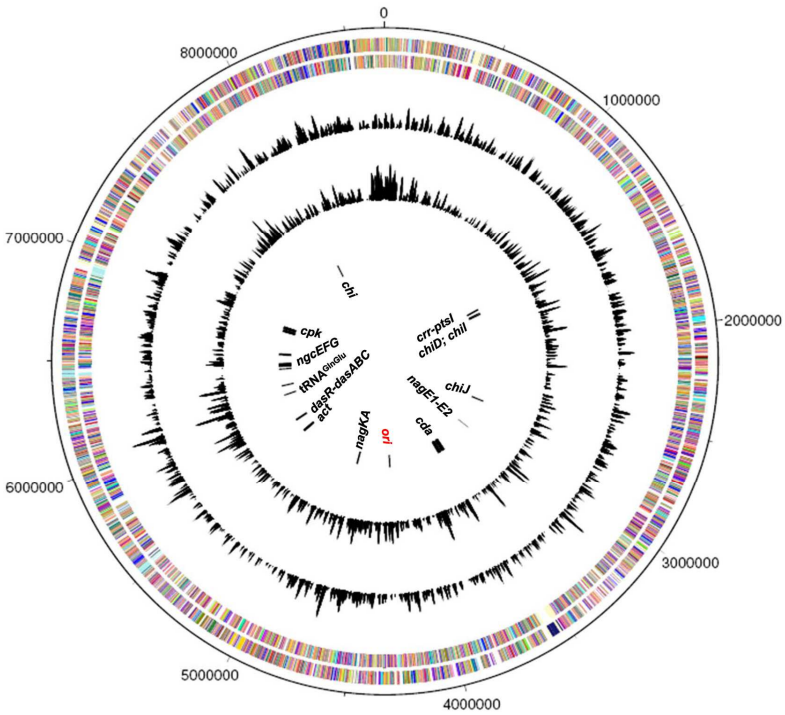

0
1000000
2000000
3000000
4000000
5000000
6000000
7000000
8000000
chi
cpk
ngcEFG
tRNA^GlnGlu
dasR-dasABC
actI
nagKA
ori
crr-ptsI
chiD; chiI
chiJ
nagE1-E2
cda

Figure 4. **Circular representation of the DasR regulon.** (A) Distribution of the DasR responsive elements (*dre)* along the *S. coelicolor* genome based on *in silico* analysis (Rigali *et al*., 2008). The heat map indicates the matrix score (Hiard *et al*., 2007) of each *dre* element relative to the consensus sequence. (B) Distribution of DasR-binding sites based on time-course ChIP-on-chip experiment. From the outside inwards, circles 1 and 2 show the position of genes transcribed in a clockwise and anticlockwise direction, respectively (for colour-coding see Table S4); circles 3 and 4 show distribution of the signal to noise ratios (log2 ratios of DasR_wt_Ab/NoAb vs DasR_mutant_Ab/NoAb) for all probes that passed filtering at 24 h and 54 h, respectively; circle 5 shows some key DasR target genes.

DISCUSSION

The DasR regulon consists of two classes of target genes

Streptomycetes usually live in an environmental context of poor nitrogen and rich carbon availability (Hodgson, 2000). In this respect, GlcNAc is a very particular energy source for these bacteria, as it can be used as both a nitrogen and a carbon source. We previously showed that DasR acts as a nutrient sensory protein that is required for development and antibiotic production under nutrient rich conditions, and that DasR binding is relieved *in vitro* by glucosamine-6P (Rigali *et al.*, 2006; Rigali *et al*., 2008). *In silico* predictions indicate widespread distribution of the *cis*-acting DasR-responsive element (*dre*) with the consensus sequence A(G/C)TGGTCTAGACCA(G/C)T. The core of the DasR regulon is formed by genes related to amino sugar metabolism, namely the *nag* and *pts* genes for GlcNAc metabolism and transport, respectively, and the *chi* regulon involved in the catabolism of polymers of GlcNAc. Since Fru-6-P, acetate and ammonia are end products of GlcNAc catabolism, the DasR core regulon is immediately connected to major metabolic pathways such as glycolysis, nucleic acids, nitrogen, and fatty acid metabolism, the tricarboxylic acid (TCA) cycle and cell-wall biosynthesis.

Functional categories of genes that are associated with *dre* elements are besides those related to *N*-acetylglucosamine metabolism and transport and polysaccharide hydrolases, also several genes for ncRNAs and tRNAs. There is an apparently perfect correlation between the presence of a *dre* motif and DasR binding *in vitro* and *in vivo*, and indeed we have never seen a predicted *dre* with a high matrix score that is not bound by DasR by itself *in vitro*. However, many binding events were identified *in vivo* that are not associated with a canonical *dre* element (Table 1). Some of these were identified as

significantly enriched probes in as many as eight biologically independent ChIP-on-chip experiments, providing a very strong case that these are indeed *bona fide* binding sites. These are referred to as Class II targets. Genome-wide analysis of the regulons of several other globally acting bacterial transcription factors show similar duality, with binding to both their cognate (consensus) target sequences as to non-canonical sites *in vivo*, even though the latter binding sites show no similarity to the consensus sequence nor are they bound *in vitro*. Examples include LexA (Wade *et al*., 2005) and FNR (Grainger *et al*., 2007) in *E. coli* and CtrA in *Caulobacter crescentus* (Laub *et al*., 2002). For *Bacillus subtilis* Spo0A, some 15% of the total number of binding sites was not bound *in vitro* (Molle *et al*., 2003). AdpA of *Streptomyces griseus* is a recent example of a master regulator in streptomycetes with a significant portion of genes without recognisable binding site (Higo *et al*., 2012). It is likely that binding by these regulators to non-canonical sites requires additional factors, such as cooperative interaction with other proteins and/or changes DNA conformation. An example of cooperative binding that is required for target recognition is that for CRP in *E. coli*, which binds to a noncanonical site in the *melAB* promoter in concert with MelR associated with adjacent target sites (Wade *et al*., 2001).

Interestingly, the conventional Class I targets are generally bound better, or - in a stochastic binding model - more frequently than then the unconventional Class II targets, as the most enriched probes in the ChIP-on-chip experiments were typically those for Class I targets. This is similar to the situation observed for LexA in *E. coli*, where unconventional targets were also not bound by LexA *in vitro*, and bound less tightly *in vivo* (Wade *et al*., 2005). A proteomics-based approach is currently being employed to elucidate how DasR binds to Class II targets *in vivo*, with particular focus on identifying possible protein partners that bind cooperaratively with DasR.

The DasR regulon and multi-level control of its target genes

The core regulon of the DasR homologue NagR in *Bacillus* also revolves around GlcNAc metabolism, but in contrast to *dre* elements, the NagR-responsive elements are found scarcely (up to three sites per genome; (Bertram *et al*., 2011)). Bioinformatics analysis suggests that similar small DasR regulons are found in other low G+C Gram-positive bacteria such as *Streptococcus* and *Listeria* (not shown). Our work shows that

streptomycetes have adopted the GlcNAc regulatory system to build a much more intricate system that controls a wide range of genes. Indeed, with transcription of more than 1200 genes significantly differentially expressed in the *dasR* null mutant of *S. coelicolor* (*i.e.* 15% of the genome) and transcription of over 400 genes altered more than two-fold, DasR qualifies as a master regulator. The regulon incorporates more than 50 transcriptional regulatory genes, suggesting an extensive secondary response regulon. Global transcriptional regulators provide an additional layer of control on top of the control mediated by specific activators and/or repressors (Martinez-Antonio & Collado-Vides, 2003). Therefore, an 'all or nothing' change in transcription as has been observed for mutants lacking specific regulatory genes such as *gylR* (glycerol regulon), *malR* (maltose regulon), *redD* (activator for prodiginine biosynthesis), or *actII*-ORF4 (activating actinorhodin biosynthesis) was not anticipated for the *dasR* regulon. The specific environmental conditions are an important parameter for the response of the regulon, and additional control systems operate, such as substrate induction for induction of sugar utilization operons, or growth phase-dependent control of antibiotic production. The additional requirement of substrate induction is exemplified by the fact that despite DasR binding to almost all of the genes of the *chi* regulon, transcription of the genes was not notably changed in *dasR* mutant cells due to the lack of chitin as inducer.

Antibiotic production is extensively controlled, so as to ensure correct growth phase-dependent expression. Taking *actII*-ORF4 as an example, this gene is controlled by some 10 transcriptional regulators. Besides DasR these include among others ActII-ORF4, AfsR, AtrA, SCO6008 and the stringent response mediated through ppGpp during growth cessation (Chakraburtty & Bibb, 1997; Floriano & Bibb, 1996; Gramajo *et al*., 1993; Uguru *et al*., 2005; reviewed in van Wezel & McDowall, 2011). DasR directly controls the pathway-specific activator genes *actII*-ORF4 and *redZ*, for Act and Red biosynthesis, respectively (Rigali *et al*., 2008 and ChIP-on-chip data presented here), as well as genes for Cpk and CDA biosynthesis, and this control is relieved by the addition of GlcNAc. This again indicates that the nutrient sensor DasR act as a typical pleiotropic regulator, which adds higher order control of gene expression to allow streptomycetes to respond to changing nutritional conditions. The small changes in the microarray experiments are due to the choice of media, necessary to allow proper comparison between wild-type and mutant, which requires similar (and synchronized) growth and

development; however, the major changes under feast (rich media) and famine (MM agar) have been well documented (Rigali *et al*., 2008).

In vivo response of DasR to the induction by GlcNAc

We previously showed that glucosamine-6P (GlcN-6P) acts as an effector of DasR *in vitro* (Rigali *et al*., 2006). This is not unexpected, as GlcN-6P stands at the crossroads of $(GlcNAc)_n$ degradation, GlcNAc transport and intracellular metabolism, glycolysis, nitrogen and lipid metabolism, as well as peptidoglycan synthesis (Altermann & Klaenhammer, 2005; Mao *et al*., 2005), and many genes of these pathways are subject to control by DasR. ChIP-on-chip analysis of MM- and R5-grown cultures before and after the addition of GlcNAc revealed the response of DasR to this inducer *in vivo*. The relief of DasR binding to its core regulon following the addition of GlcNAc to cultures provides *in vivo* evidence that the DasR regulon is indeed triggered by GlcNAc-derived metabolites. In terms of antibiotic production, the same relief of binding was seen for the pathway-specific activator genes *redD* and *redZ* for prodiginine production, and *actII*-ORF4 and *atrA* for actinorhodin production, as well as biosynthetic genes for the cryptic polyketide Cpk and the calcium-dependent antibiotic CDA. This suggests that the activation of antibiotic production by GlcNAc is directly mediated through DasR in *S. coelicolor*.

DasR and translational control

An important control system that has hitherto not been identified in streptomycetes but which must almost inevitably be in place is that of the control of ribosomal proteins and stable RNA operons (specifying tRNA and rRNA). In *E. coli*, stable RNA promoters are regulated by the trans-acting proteins FIS (Factor for Inversion Stimulation), which activates transcription during logarithmic growth (Nilsson *et al*., 1990) and LRP which - assisted by H-NS - forms a repressive nucleoprotein structure (Pul *et al*., 2007). The mechanisms for ribosome feedback regulation and growth rate-dependent control of stable RNA synthesis in streptomycetes are poorly understood. The ChIP-on-chip data demonstrate binding of DasR to at least five tRNA genes or operons when *S. coelicolor* was grown in liquid R5 cultures. This binding was significantly enhanced after addition

of GlcNAc. Strong binding of DasR was observed to a *dre* element with one of the highest matrix scores upstream of the third gene of a five-members tRNA operon consisting of three identical genes specifying tRNA-Glu for the CUC codon and two specifying tRNA-Gln for the CUG codon. By leaving the first two tRNA genes (for tRNA-Gln and tRNA-Glu, respectively) untouched but controlling the expression of the other (identical) copies, DasR likely controls the abundance of these tRNAs. Many independent ChIP-on-chip experiments revealed DasR binding to many probes inside the rRNA (*rrn*) operons, suggesting that DasR may also control the expression of rRNAs (Table 1 and Table S3).

Besides stable RNAs, DasR also controls the expression of non-coding RNAs. In at least two independent ChIP-on-chip experiments DasR binding was associated with the upstream regions of the ncRNAs scr3092 and scr5239. The latter is preceded by a highly conserved *dre* element (Table 1), and deletion of *dasR* results in a four-fold increase of expression of scr5239 (not shown). The current generation of high-density *S. coelicolor* microarrays is primarily directed at identifying mRNAs, but better insight in the location of ncRNAs will soon be obtained from RNA seq experiments, so that we can obtain better insight into how many are controlled by DasR, and how.

Microarray analysis further revealed that the expression of nearly all the ribosomal protein genes was affected in the *dasR* mutant. Since with the exception of *rpmG* these genes were not identified as direct targets, these changes in expression either reflect control by a DasR-dependent regulator, or general differences in growth between the parental strain and its *dasR* null mutant. The possible role of DasR in the control of stable RNA and perhaps also ribosomal protein genes requires further investigation.

Connection of major developmental regulons

The phenotype of the *dasR* null mutant as well as extracellular complementation assays strongly suggests it is blocked in a very early stage of the life cycle, and *dasR* is probably one of the *bld* mutants acting earliest in the life cycle (Rigali *et al*., 2006; Rigali *et al*., 2008). Our data indicate that DasR exerts this early developmental control by targeting the expression of a number of key regulatory genes for the control of early developmental processes, including *bldKc, bldM, bldN, wblA, wblC* and *crp*. Of these,

bldKc, *bldN* and *wblC* were identified as direct targets by ChIP-on-chip analysis, forming Class II targets as none of them is preceded by a *dre* element. In turn, *bldN* gene product σ^{bldN} is required for transcription of the *chp*, *rdl* and *ramS* genes for the surfactants chaplins, rodlins and SapB, respectively (Bibb *et al*., 2012). The increase in *bldN* transcription corresponds to an increase in *chp*, *rdl* and *ramS* expression during sporulation, but at earlier time points all genes were strongly down-regulated. The DasR-controlled regulatory gene responsible for this phenomenon has not yet been identified.

The *bldD* gene also emerged as a possible direct target for DasR, with binding to the *bldD* promoter region verified independently in samples obtained after 54 h of growth on MM agar plates as well as from liquid-grown R5 cultures, although no *dre* element could be identified upstream of *bldD*. The transcription of *bldD* was not changed in the *dasR* deletion mutant, but the sporulation block of most *bld* mutants is by-passed on minimal media with mannitol, suggesting their function is limited under these conditions. Considering the scale at which DasR and BldD control gene expression in streptomycetes, there is surprisingly little overlap between the two regulatory networks. Regulators of later sporulation events do not feature among the direct DasR targets, while, conversely, the primary BldD regulon contains a plethora of such developmental regulatory genes (den Hengst *et al*., 2010), suggesting that these two proteins act in different time domains. The level of cross-talk or hierarchy between these two important regulatory networks deserves closer investigation.

In conclusion, our works establishes DasR as a master regulator that links nutritional control to antibiotic production and development (Fig. 5), and the exogenous addition of GlcNAc relieves binding by DasR to many of its binding sites. This supports the notion that accumulation of GlcNAc around *Streptomyces* colonies affects development via the direct interference with binding of the nutrient sensor DasR. The canonical binding of DasR to Class I targets is observed both *in vivo* and *in vitro*, while the unconventional binding to Class II targets presumably requires cooperative binding with another protein, which has hitherto not been identified. Alternatively, these sites may require conformational changes in the DNA, or even a combination of the two. DasR mediates a higher order level of control, and relieve or induction of binding to its target sites does not necessarily result in an immediate change in gene expression. For that, the environmental conditions are a very important determining factor - such as the presence

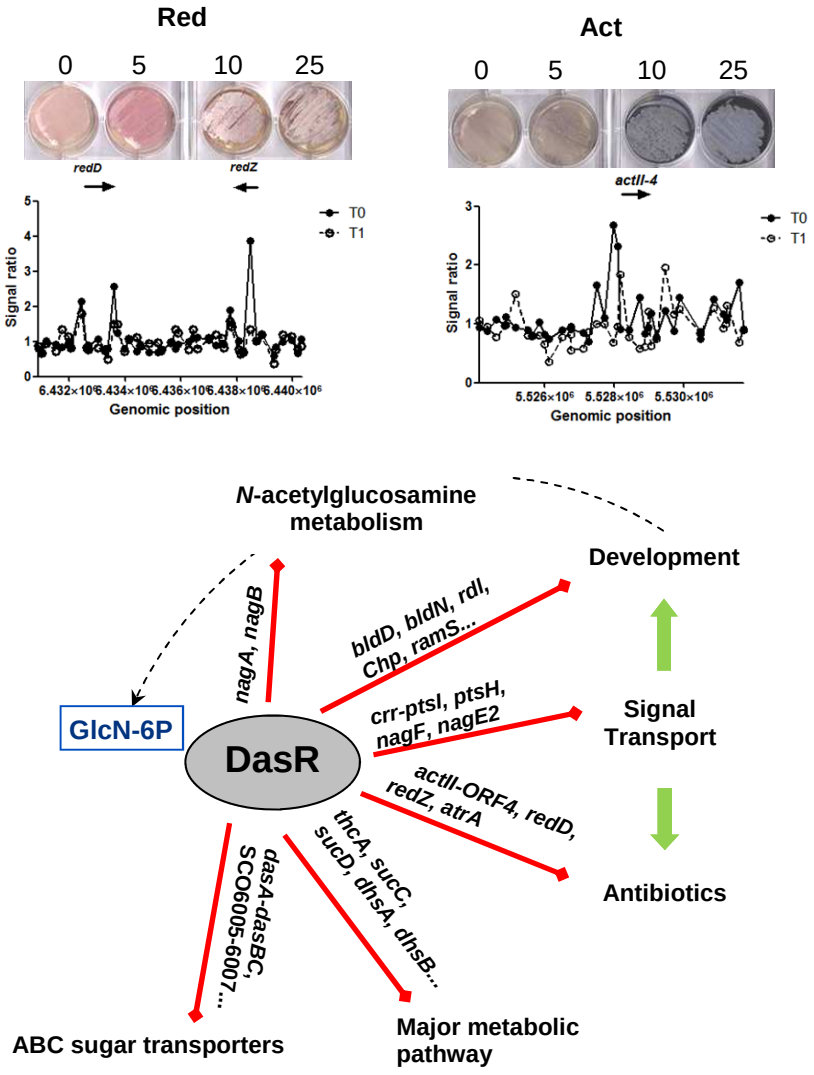


Figure 5. Model of the DasR regulon and its induction by *N*-acetylglucosamine. Top: induction of development (grey pigment) and Red (left; red pigment) or Act (right; blue pigment) production on MM agar plates with increasing concentrations of *N*-acetylglucosamine; the *S. coelicolor actII*-ORF4 (M510) and *redD* (M511) mutants (Floriano & Bibb, 1996) were used to demonstrate the production of only the red-pigmented undecylprodigiosin or the blue-pigmented actinorhodin, respectively. **Middle,** relief of DasR binding *in vivo* after the addition of GlcNAc, as shown by ChIP-on-chip (refer to Fig. 3). T_0 (closed circles) and T_1 (open circles), time points before and 30 min after the addition of 25 mM GlcNAc; respectively. **Bottom,** summary of the DasR-controlled processes and major components of its regulon. Red lines indicate direct binding and transcriptional repression of the targeted genes by DasR. Depression of the DasR regulon in response to the accumulation of GlcN-6P results in activation (green lines) of among others development and antibiotic production.

of chitin for induction of *chi* gene expression - and further experimental data need to be obtained from genome-wide studies under different growth conditions. For example, in approaches to wake up sleeping antibiotic genes, relieve of DasR may well be an important step, but better understanding of the triggers that must be in place to effectively activate their biosynthetic gene clusters should aid in their discovery and exploitation.

Acknowledgements

This work was supported by a VICI grant of the Netherland technology Foundation STW to GPvW.