Molecular characterization of copper-dependent enzymes involved in Streptomyces morphology

Petrus, Maria Louise Catharina

6

Abstract

Extracellular polysaccharides are produced by organisms from all kingdoms of life and play crucial roles in their biology. The filamentous bacterium *Streptomyces lividans* produces an extracellular glycan at hyphal tips by the action of the cellulose synthase-like protein CslA. The absence of CslA has a dramatic effect on morphology, as it blocks the formation of pellets in liquid environments and reproductive aerial structures on solid substrates. Despite its importance in morphogenesis, the glycan produced by this enzyme has not been characterized. Using a bioinformatics approach we show that CslA is a processive glycosyltransferase with a three dimensional structure comparable to the cellulose synthase BcsA of *Rhodobacter spaeroides.* However, differences are present in specific motifs involved in substrate and product binding. Subsequent chemical characterization of the polysaccharide was severely hampered by the relative abundance of peptidoglycan. To overcome this hurdle, we developed a cell wall-deficient synthesis platform for glycan characterization. As a proof-of-principle, we used this system to determine the composition of the CslA polymer and show that it may consist of glucose and/or *N*-acetylgalactosamine. We anticipate that this innovative production platform will facilitate characterization of glycans that are only produced in small amounts and for which purification is difficult.

Marloes L.C. Petrus, Sander S. van Leeuwen, Karina Ramijan, Karthick B.S.S. Gupta, Wim Jesse, Erik Vijgenboom, Gilles P. van Wezel, Young Choi, Lubbert Dijkhuizen, Dennis Claessen

Use of an innovative peptidoglycan-independent platform for characterization of the glycan produced by CslA

Introduction

Cellulose and chitin are the most abundant organic compounds in nature. They are key components of the cell walls of plants, algae, fungi or yeast, but also the exoskeleton of arthropods, where their main function is to provide structural rigidity. Cellulose is also produced by bacteria, where it contributes to the formation of sessile, multicellular growth forms called biofilms (Saxena *et al.*, 1990; Ross *et al.*, 1991; Römling, 2002; Jahn *et al.*, 2011; Römling and Galperin, 2015). Polysaccharide synthesis is catalyzed by glycosyltransferases (GTs) that exhibit a strong specificity for both the donor and the acceptor molecules (Breton *et al.*, 2006; Lairson *et al.*, 2008). For instance, cellulose synthases use UDP-activated glucose residues as donor substrate, while chitin synthases use UDP-activated *N*-acetylglucosamine. Both synthases incorporate these sugars into a growing glycan chain by catalyzing the formation of β-(1,4)-glycosidic bonds. Due to their similar reaction mechanism both synthases are classified into the broad group of family 2 GTs, according to the Carbohydrate Active enZyme database (CAZy; Lombard *et al.*, 2014).

In enteric bacteria cellulose biosynthesis is performed by a protein complex of three subunits: BcsA, BcsB, and BcsC (Ross *et al.*, 1991). BcsA is an integral inner membrane protein with an intracellular catalytic domain, BcsB is a membrane-associated periplasmic protein, while BcsC is predicted to form a pore in the outer membrane. Only BcsA and BcsB are essential for cellulose synthesis *in vitro,* however, also the action of BcsC and the cellulase BcsZ are required for synthesis *in vivo* (Römling, 2002; Römling and Galperin, 2015). In 2013 the catalytically-active BcsA-BcsB complex of *Rhodobacter sphaeroides* was crystalized, which also contained the polysaccharide that was being synthesized (Morgan *et al.*, 2013). The structure revealed that BcsA contains eight transmembrane (TM) helices, of which TM3

to TM8 form a narrow channel for transport of this polysaccharide over the plasma membrane (Fig. 1A). The active site of BcsA is located between TM helices 4 and 5, and adopts a canonical glycosyltransferase A (GT-A) fold (Charnock and Davies, 1999; Breton *et al.*, 2006). The intracellular C-terminus of the protein contains a PilZ domain that can bind cyclic-di-GMP, thereby inducing a conformational change in BcsA leading to activation of the complex. BcsB is largely located on the periplasmic side and interacts via carbohydrate binding domains with the polysaccharide.

The crystal structure of BcsA provides important insight in the role of the highly conserved residues of the characteristic D,D,D,Q(Q/R)XRW signature present in all processive βglycosyltransferases (Saxena *et al.*, 1995; Morgan *et al.*, 2013; Gloster, 2014). In BcsA, the first two variably spaced aspartic acid residues of the signature (Asp179 and Asp246), located in the DDG and DXD motifs, coordinate UDP-glucose in the active site, while the

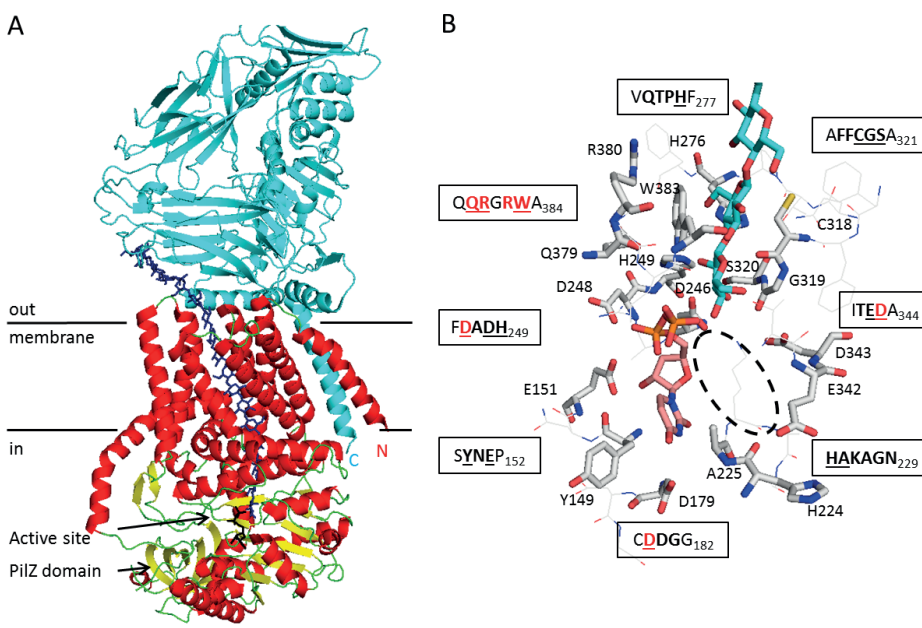

Figure 1. *Architecture of the BcsA-B complex from Rhodobacter sphaeroides (Morgan et al., 2013).* ***A****) Crystal structure of the BscA-BscB complex, with BcsB colored in cyan. BcsA helices are shown in red, while loops and strands are shown in green and yellow, respectively. The translocating glucan is shown in dark blue and the UDP-moiety is shown in black in the active site.* ***B****) Active site residues of BcsA are engaged in coordinating the UDP moiety of UDP-glucose and the terminal disaccharide of the translocating glucan. The organic molecules are represented in sticks with UDP in orange and the glucan in cyan. Conserved amino acids are represented in grey colored sticks and belong to sequence motifs shown in single letter code. Conserved residues are highlighted in bold and the depicted residues are underlined. The residues of the glycosyltransferase signature are shown in red letters. The likely position of the donor Glc is indicated by a dashed ellipsoid. Images were recreated in PyMol based on figures from Morgan et al. (2013).*

third aspartic acid (Asp343), located in the TED motif, probably acts as the catalytic base (Morgan *et al.*, 2013; Fig. 1B). The Q(Q/R)XRW part of the signature is located at the interface between the cytosol and the membrane. There, Trp383 interacts with the penultimate sugar moiety of the acceptor glycan.

In addition to the processive glycosyltransferase signature also the role of certain cellulose synthase-specific motifs is shown by the crystal structure. The amino acid residues of the FFCGS motif are engaged in coordinating the ultimate glucose moiety of the growing chain (Fig. 1B). The QTPH cellulose synthase-motif is also involved in positioning the translocating glycan as it is within hydrogen bonding distance to the penultimate glucose via the histidine, which in turn is positioned properly by the proline residue contained in this motif. Residues in the HXKAG motif are involved in coordinating UDP-glucose in the binding pocket together with the glutamate of the TED motif (Fig. 1B).

The filamentous soil bacterium *Streptomyces lividans* contains a family 2 GT called $CslA_{SL}$, for cellulose synthase-like protein. The $CslA_{SL}$ protein makes a glycan at hyphal tips with multiple roles during growth and development; it is required for the formation of aerial hyphae on solid media (Xu *et al.*, 2008), and also for the formation of large mycelial aggregates, called pellets, in liquid-grown cultures (Xu *et al.*, 2008; Chaplin *et al.*, 2015). Furthermore, the CslAderived glycan is involved in the attachment of hyphae to hydrophobic surfaces (de Jong *et al.*, 2009). Despite its importance, the true identity of the glycan produced by $CslA_{SL}$ remains to be discovered. Preliminary experimental evidence suggested that cellulose-like polymers are formed at apical sites. This is based on 1) the $CslA_{SL}$-dependent hyphal tip staining with calcofluor white (CFW), which stains β-(1,3)- and β-(1,4)-coupled glycans including cellulose and chitin, and 2) the abolishment in hyphal attachment of the wild-type strain when grown in the presence of cellulase, which mimics the phenotype of a $cslA_{SL}$ mutant (Xu *et al.*, 2008; de Jong, Wösten, *et al.*, 2009). However, the purity of enzyme preparations and also the specificity of the enzymes themselves is often questionable (Schiavone *et al.*, 2014).

Purification and characterization of the glycan produced by $CslA_{SL}$ is hampered by the relative little amount of this polymer that appears to be produced compared to peptidoglycan (PG), which is the major constituent of the *Streptomyces* cell wall. To circumvent these problems, we here describe an innovative PG-independent synthesis platform for characterization of unknown glycans, with the polymer produced by CslA as a proof-of-concept. Preliminary data using this system indicated that the CslA polymer could be composed of glucose and/or *N*-acetylgalactosamine. We anticipate that our system may be applied for the synthesis and purification of glycans from other bacterial sources that in their endogenous hosts are only produced in limited amount.

Results

Gene synteny correlates cslA$_{SL}$ to polysaccharide biosynthetic genes

The genetic organization around *cslA*$_{SL}$ is different from the *bcsABZC* cellulose biosynthesis operon in enteric bacteria (Römling, 2002), since *cslA*$_{SL}$ (SLI_3187) is organized in a cluster with *glxA*$_{SL}$ (SLI_3188) for a radical copper oxidase and *cslZ* (SLI_3189) for a β-(1,4)-endoglucanase (Liman *et al.*, 2013). In terms of gene synteny, conserved genes are found in close proximity to this gene cluster in nearly all streptomycetes (Fig. 2A): SLI_3182 encodes a copper-dependent lytic polysaccharide monooxygenase (LPMO), which typically play important roles in the degradation of polysaccharides (Nakagawa *et al.*, 2015); SLI_3183 encodes a protein with a flotillin domain, which is implicated in localizing large multiprotein complexes to the membrane; SLI_3184 encodes a protein with a peptidoglycan-binding domain and an uncharacterized NLPC_P60 domain that is found in some hydrolases; SLI_3192 encodes for sortase E (SrtE), an enzyme that covalently couples proteins to the cell wall, with SLI_3193 encoding such a putative sortase E substrate; and SLI_3194 encodes NagD, a possible ribonucleotide monophosphatase, which might play a role in recycling UDP when liberated by the action of glycosyltransferases on UDP-activated sugar moieties (Tremblay *et al.*, 2006). Thus, several genes in this genetic location are involved in polysaccharide synthesis and/or modification.

Bioinformatics analysis and structural modeling of CslA$_{SL}$

cslA$_{SL}$ was annotated as a cellulose synthase-like gene based on similarity to the catalytic subunits of cellulose synthases from other bacteria and plants (Xu *et al.*, 2008). Careful bioinformatics analyses, based on new prediction programs and increased understanding of these synthases, indicated that CslA$_{SL}$ is an integral membrane protein with six transmembrane (TM) helices (Fig. 2B), in contrast to the earlier prediction of seven TM helices (Xu *et al.*, 2008). Both the N- and the C-terminus of the protein are predicted to reside in the cytoplasm. The first two TM helices are followed by a large cytoplasmic region (aa 143 – 425) that contains a CESA_CelA_like domain (aa 168 – 406), which is also found in the known cellulose synthases CslA$_{GX}$ of *Gluconacetobacter xylinus,* CelA of *Agrobacterium tumefaciens* and BcsA of *R. sphaeroides* (Saxena *et al.*, 1990; Matthysse *et al.*, 1995; Morgan *et al.*, 2013). The other four TM helices follow this cytoplasmic domain and are near the Cterminus of the protein.

Alignment of the amino acid sequences of the cytoplasmic domain of *S. lividans* CslA$_{SL}$ against those of the above-mentioned cellulose synthases indicates that the D,D,D,Q(Q/R)XRW signature is present (Xu *et al.*, 2008; Fig. 2C), confirming CslA$_{SL}$ as a

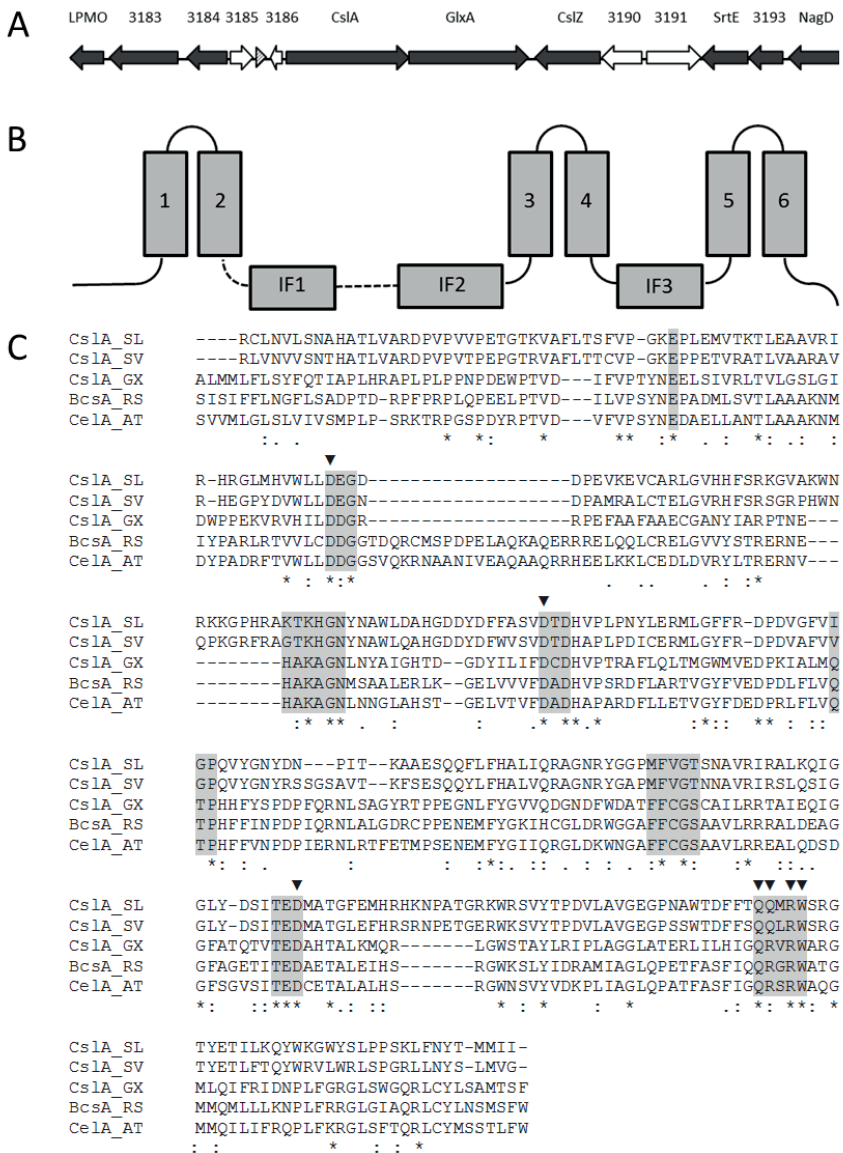

Figure 2. *Bioinformatics analysis of CslA$_{SL}$.* ***A)*** *Gene organization around cslA$_{SL}$ in S. lividans. The filled arrows represent genes that are conserved in this region in nearly all streptomycetes. The open arrows represent genes that are not conserved, while the dashed arrowhead represents a tRNA locus. Numbers correspond to the SLI number as given by Cruz-Morales et al. (2013).* ***B)*** *Schematic representation of the six transmembrane helices (1-6) and three interfacial helices (IF1-3) of CslA$_{SL}$ as predicted by TMHMM 2.0 and Phyre. The catalytic site is conserved in the cytoplasmic region between transmembrane helices 2 and 4.* ***C)*** *Protein sequence alignment of the catalytic regions of CslA$_{SL}$ of S. lividans (EOY47900) with CslA$_{SV}$ of S. viridifaciens, CslA$_{GX}$ of Gluconacetobacter xylinus (CAA38487.1), BcsA of Rhodobacter spaeroides (WP_041669585.1) and CelA of Agrobacterium tumefaciens (NP_533806.1) with Clustal Omega. Conserved motifs are highlighted in grey. The residues of the D,D,D,Q(Q/R)XRW signature are indicated with arrow heads above the alignment.*

processive glycosyltransferase. However, some other motifs uniquely found in cellulose synthases are less well conserved (*vide infra*). We then used the Phyre2 server (Kelley *et al.*, 2015) to predict the 3D structure of $CslA_{SL}$. The highest scoring structural template used for 3D conformation modeling of $CslA_{SL}$ is the crystal structure of the cellulose synthase BcsA from *R. spaeroides*, which has a 24% identity score in the aligned region. Modeling of 529 residues of $CslA_{SL}$ (i.e. residues 79-608) leads to the structure prediction shown in Fig. 3. Comparison of this model with the 3D-structure of BcsA (Fig. 1) indicates that the six TM helices of $CslA_{SL}$ form the glycan-translocating channel, while the cytoplasmic CESA_CelA_ like domain in $CslA_{SL}$ has a similar fold as the comparable domain in BscA. However, BcsA has two additional N-terminal TM helices, which are not present in $CslA_{SL}$. One of these two helices may interact with the C-terminal TM helix of BcsB, which is missing in *S. lividans.* Also absent in $CslA_{SL}$ is the C-terminal PilZ domain of BcsA, which is involved in activation of the protein via interaction with c-di-GMP (Morgan *et al.*, 2014).

Detailed analysis of the active site of $CslA_{SL}$ indicates that the organization of the D,D,D,Q(Q/R)XRW signature is very similar to the organization in BcsA (Fig. 3B). However, the hydrophobic alanine underlined in the HAKAG motif that flanks the binding pocket for UDP-Glc in BcsA is substituted by a larger, polar threonine in $CslA_{SL}$ (KTKHG). Changes in the

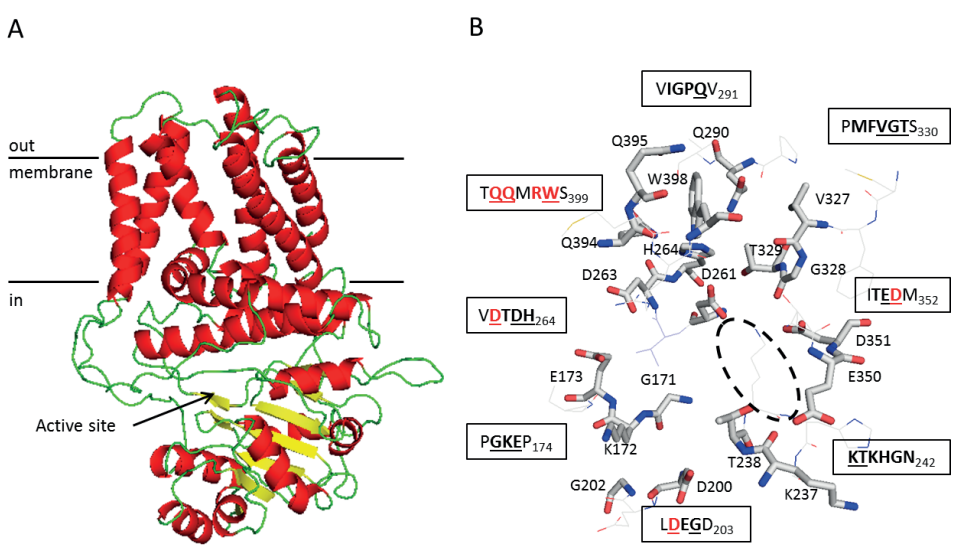

Figure 3. *Computational model and the active site interacting residues of $CslA_{SL}$ from S. lividans.* ***A)*** *Cartoon representation with $CslA_{SL}$ colored according to its predicted secondary structure with helices in red, loops in green and strands in yellow. Horizontal bars indicate the membrane boundaries.* ***B)*** *Conserved amino acids of the predicted active site of $CslA_{SL}$ are represented in grey colored sticks and belong to sequence motifs shown in single letter code. Conserved residues are highlighted in bold and the depicted residues are underlined. The residues of the glycosyltransferase signature are shown in red letters. The likely position of the donor sugar is indicated by a dashed ellipsoid. Images were prepared in PyMol.*

conserved active site residues YNE possibly affect interactions with the UDP-site of the activated sugar. Also interactions with the acceptor glycan chain appear to be different due to differences in the QTPH/IGPQ motif, where the polar Gln290 residue is distinct from the positively charged His276 residue, and the FFCGS/MFVGT motif, where Thr329 is larger than Ser320. Note hereby that $CslA_{SL}$ is modeled along the backbone of BcsA and that changes in the backbone orientation might occur that are not visible in this predicted 3D-structure.

Taken together, $CslA_{SL}$ shows the characteristics of a cell membrane-embedded processive glycosyltransferase with an overall fold very similar to that of BcsA. However, the changes in several residues near the catalytic site of $CslA_{SL}$ may lead to different substrate specificity in $CslA_{SL}$.

Cellulase or chitinase treatment abolish hyphal attachment

Attachment of *S. coelicolor* hyphae to surfaces is dependent on the activity of CslA (de Jong, *et al.*, 2009). In agreement, deletion of *cslA*$_{SL}$ in *S. lividans* reduces attachment to 53 ± 5% of the wild type, while deletion of the functionally-related *glxA*$_{SL}$ gene reduces hyphal attachment to 32 ± 6% (Fig. 4A). Considering that $CslA_{SL}$ is a processive glycosyltransferase, we then assessed attachment of the wild-type strain in the presence of increasing amounts of cellulase and chitinase. Growth in the presence of cellulase abolished attachment, consistent with earlier results in *S. coelicolor* (Fig. 4B). Notably, growth in the presence of increasing amounts of chitinase lead to a similar loss of adhesion (Fig. 4B). In contrast, O-glycosidase, which catalyzes the removal of O-linked disaccharides from glycoproteins, had no effect on attachment (data not shown).

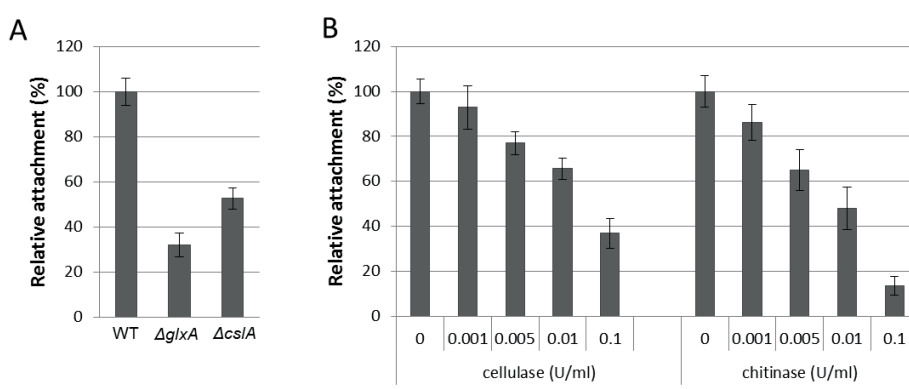

Figure 4. *Attachment of mycelium to hydrophobic surfaces.* ***A****) S. lividans cslA*$_{SL}$ *and glxa*$_{SL}$ *deletion mutants show reduced attachment in liquid standing mNNMP cultures.* ***B****) Increasing amounts of cellulase or chitinase reduce the attachment of hyphae to hydrophobic surfaces. Error bars indicate standard error of the mean measured from thirteen biological replicates in A and eight biological replicates in B.*

Peptidoglycan hampers the characterization of the CslA$_{SL}$-dependent glycan

Cell walls of the parental strain *S. lividans* 1326 and its *cslA*$_{SL}$ deletion derivative were isolated and ground with potassium bromide for analysis with Fourier transform infrared spectroscopy (FTIR), which is commonly used for the identification of carbohydrates (Kumirska *et al.*, 2010). Both samples showed a similar spectrum with strong absorbance bands of amide I (1650 cm^{-1}) and amide II (1550 cm^{-1}; Fig. 5) characteristic of the presence of *N*-acetyl groups. Given the presence of these signals in both strains, we anticipate that they result from *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) residues in the peptidoglycan.

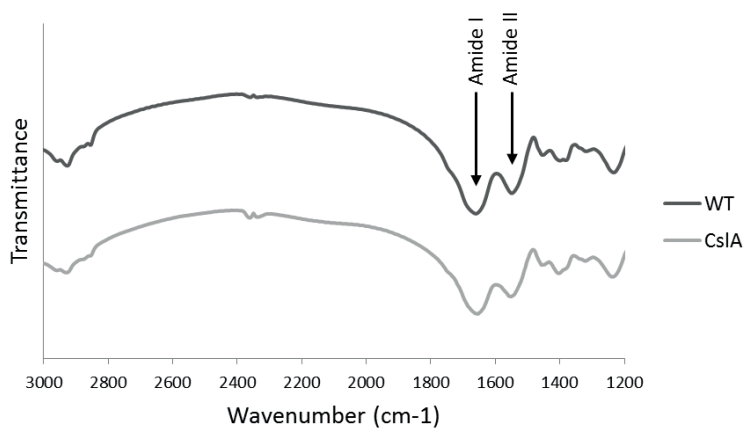

Figure 5. *FTIR spectra of cell wall fractions of S. lividans WT and ΔcslA*$_{SL}$ *are identical. The specific amide I (~1650 cm*$^{-1}$*) and amide II (~1550 cm*$^{-1}$*) absorbance bands are indicated by arrows.*

We also analysed cell walls using ^{13}C CPMAS solid-state NMR, which provides information on composition and glycan structure (Fig. 6). The spectra of *S. lividans* 1326 (A) and Δ*cslA*$_{SL}$ (B) showed almost the same peak pattern. To understand the origin of these rather complex spectra we used the same approach to analyse peptidoglycan from *Bacillus subtilis* (C), purified chitin (D) or cellulose (E). The spectrum of cellulose shows the peaks of sugar carbon atoms, while in the spectrum of chitin two additional characteristic peaks resulting from the *N*-acetyl groups of GlcNAc were seen: The very sharp signal at 22.6 ppm is from the methyl carbon atoms, and the peak at 173.1 ppm is from the carbonyl carbon atoms. In the spectrum of peptidoglycan the peaks are broader than in the spectra of cellulose and chitin because peptidoglycan is not as crystalline. In addition to the sugar carbons of GlcNAc and MurNAc in the peptidoglycan backbone, also the carbon atoms of the peptide bridges contribute to the peaks in the spectrum, where they overlap with the before mentioned carbonyl carbon, sugar carbon and methyl carbon peaks (Bougault *et al.*, 2012). The spectra

of *S. lividans* 1326 (A) and Δ*cslA*$_{SL}$ (B) most likely result from a similar combination of peaks derived from cell wall peptidoglycan, including peaks resulting from sugar carbons, *N*-acetyl groups and peptide carbons. As both spectra show a virtually identical peak pattern it seems that the amount of glycan produced by CslA is only small in respect to the abundant peptidoglycan.

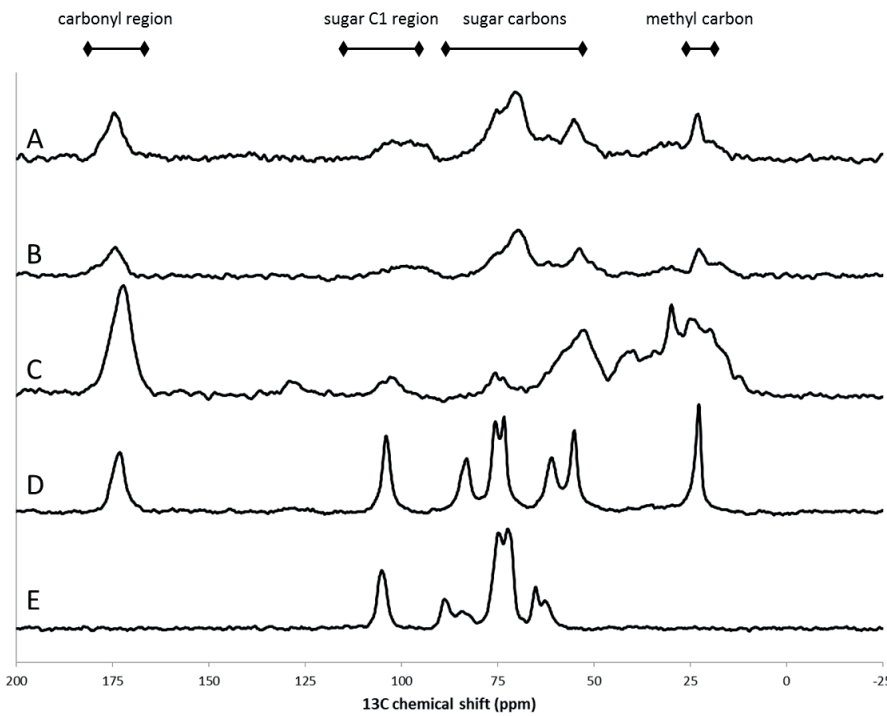

Figure 6. *Solid-state ^{13}C CP-MAS NMR spectra of cell wall fractions of S. lividans WT* ***(A)*** *and Δ*cslA*$_{SL}$ ***(B)****, peptidoglycan from Bacillus subtilis* ***(C)****, chitin* ***(D)*** *and cellulose* ***(E)****. Main regions that can be detected in the NMR spectra of peptidoglycan or polysaccharide are indicated above the graph, with horizontal lines indication the corresponding region.*

To analyse the monosaccharide composition of cell walls in more detail, we methanolyzed cell wall fractions of the *S. lividans* wild-type strain and compared these to fractions obtained from strains lacking *cslA*$_{SL}$ or *glxA*$_{SL}$. The biological replicates showed considerable differences both in monosaccharide composition and in the percentage of the total sugars content, which varied from 18 to 79 % in the *glxA*$_{SL}$ mutant (Fig. 7). The analysis showed that the cell walls of the *cslA*$_{SL}$ and *glxA*$_{SL}$ deletion mutants mainly consist of GlcNAc or sugars that are converted into this sugar during the methanolysis process (such as MurNAc). The wild-type samples contained 36 % and 23 % glucose respectively, unlike the mutant strains where glucose was hardly detectable (Fig. 7). This indicates that the polysaccharide composition of the cell wall is changed in the absence either *cslA*$_{SL}$ or *glxA*$_{SL}$.

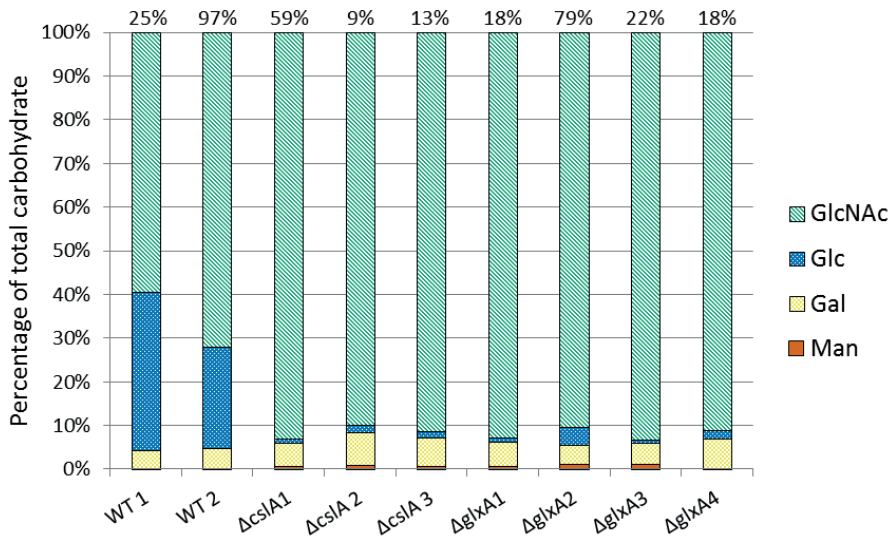

Figure 7. *Carbohydrate monomer composition of cell walls from TSBS cultures of S. lividans WT, Δ$cslA_{SL}$ and Δ$glxA_{SL}$, as detected after acid hydrolysis and GLC-MS quantification of pentose and hexose sugars (colored as indicated in figure legend). Values are represented as a mole-based percentage of total recovered carbohydrate. Percentages above each bar represent the fraction of carbohydrate as a percentage of the total weight of the sample. The abbreviations used are: GlcNAc, N-acetylglucosamine; Glc, glucose; Gal, galactose; Man, mannose.*

Use of cell wall deficient L-forms for glycan identification

Our data indicate that the presence of PG complicates the analysis of the glycan produced by $CslA_{SL}$. We therefore considered developing a platform in which the synthesis of PG is abolished. Growth without PG results in the formation of so-called L-forms, which have been created in several actinomycetes, including *Streptomyces viridifaciens* (Innes and Allan, 2001). Based on phylogenetic analysis and genome comparisons this species has recently been shown to belong to the genus of *Kitasatospora* (Girard *et al.*, 2014), which is closely related to the genus *Streptomyces* in the family Streptomycetaceae (Labeda *et al.*, 2012). Bioinformatic analysis shows that the *$cslA_{SL}$-$glxA_{SL}$* gene cluster is conserved in *S. viridifaciens*, encoding proteins that share 53% and 52% aa identity to $CslA_{SL}$ and $GlxA_{SL}$, respectively. Amino acid alignment of the catalytic domain of $CslA_{SV}$ shows that the glycosyltransferase-specific signature is completely conserved (Fig. 2B). Moreover, all amino acids of $CslA_{SL}$ that are predicted to interact with the UDP-sugar moiety or the translocating glycan are conserved in $CslA_{SV}$ (Fig. 2B, Fig. 3B). The genetic organization around the *$cslA_{SV}glxA_{SV}$* cluster is similar to the organization in *S. lividans*, showing the neighborhood of NagD and a gene encoding an endoglucanase (Fig. 8A, Table 2). Together, this gives a strong indication that these homologues fulfill the same biological role.

To test the biological role of *cslA*$_{SV}$ or *glxA*$_{SV}$, deletion mutants were created in *S. viridifaciens* and compared to the corresponding *S. lividans* mutants. Both the *S. viridifaciens* and *S. lividans* wild-type strains form a robust aerial mycelium and produce spores after four days of growth on MYM plates (Fig. 8B-C). Deletion of *glxA* blocks aerial growth in both organisms, while aerial mycelium formation and sporulation is greatly reduced in the *cslA* deletion mutants. Morphology of the mutants was also studied in liquid-grown cultures. In contrast to *S. lividans*, *S. viridifaciens* does not form dense pellets in liquid-grown cultures

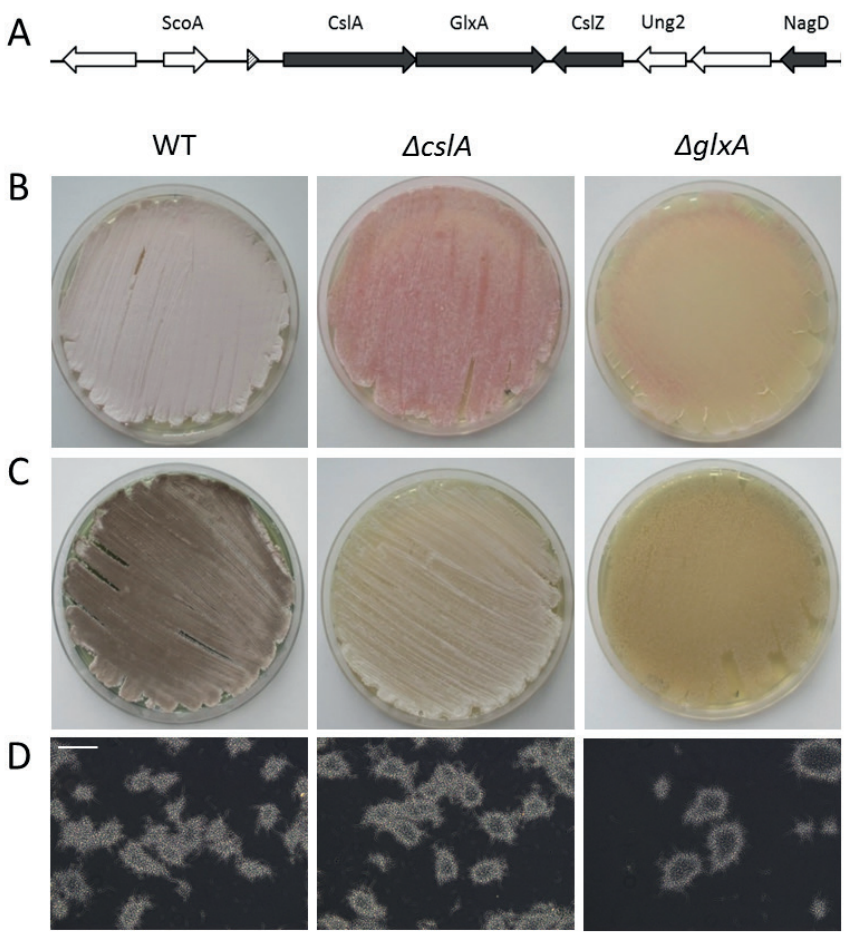

Figure 8. *The S. viridifaciens cslA*$_{SV}$*-glxA*$_{SV}$ *gene cluster.* ***A)*** *Gene organization around cslA*$_{SV}$*. The filled arrows represent genes that are conserved in this region in nearly all streptomycetes. The open arrows represent genes that are not conserved, while the dashed arrowhead represents a tRNA locus.* ***B-C)*** *Phenotypes WT, ΔcslA and ΔglxA strains of S. lividans* ***(B)*** *and S. viridifaciens* ***(C)*** *after 4 days of growth on MYM agar plates. Note that the glxA mutants do not produce any aerial mycelium, while the cslA mutants are severely delayed in aerial mycelium and spore formation.* ***D)*** *Morphology of S. viridifaciens WT, ΔcslA*$_{SV}$ *and ΔglxA*$_{SV}$ *strains after 1 day of growth in liquid shaken TSBS. Scale bar is 100 μm.*

(Fig. 8D). As a consequence, deletion of *cslA*$_{SV}$ or *glxA*$_{SV}$ had no effect on mycelial architecture, unlike the absence of the corresponding genes in *S. lividans,* which prevents the formation of pellets (Xu *et al.*, 2008; Chaplin *et al.*, 2015). Taken together, the strong similarity of the CslA$_{SL}$ and CslA$_{SV}$ active sites, and the equivalent morphological defects of the *cslA* mutants infer that both synthases have a comparable role and produce a similar glycan.

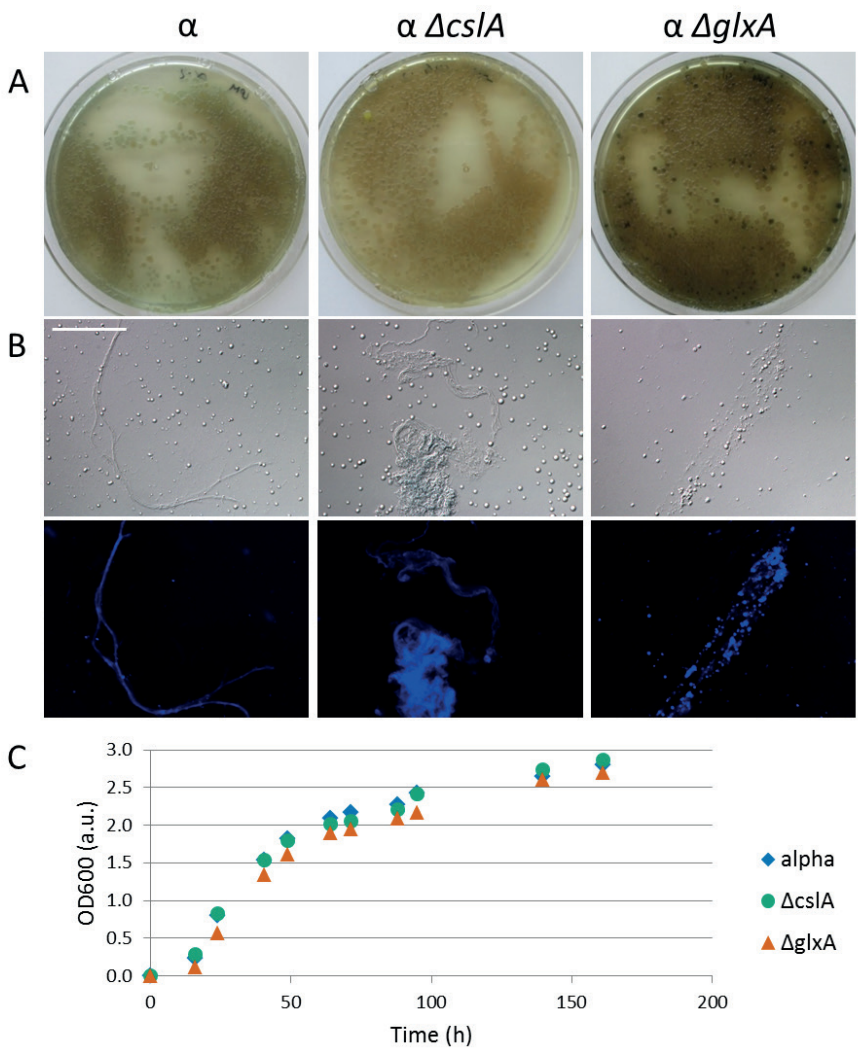

Figure 9. *Phenotypic analysis of L-form ΔcslA*$_{SV}$ *and ΔglxA*$_{SV}$ *mutants.* ***A)*** *Morphology of S. viridifaciens L-form strains alpha (parental strain), ΔcslA*$_{SV}$ *and ΔglxA*$_{SV}$ *after 3 days of growth on LPM agar plates.* ***B)*** *L-forms and produced fibers stained with CFW from a 7 days old LPB culture. Upper row are the brightfield images, lower row is CFW fluorescence signal. Scale bar is 100 µm. C) Growth curves of L-forms strains in LPB is measured by OD600 and corrected for the absorption of the green pigment.*

In order to study glycan production in the L-form background, which supposedly lacks PG, we constructed a series of control strains in which the *cslA*$_{SV}$ and *glxA*$_{SV}$ genes were replaced by an apramycin resistance cassette (see M&M for details). Mutants were grown on LPM agar plates and compared to the parental strain. Growth of the wild-type and mutant L-form strains resulted in the formation of colonies with a smooth, glossy surface that were similar in appearance (Fig. 9A). Also, the morphology and growth rate in liquid LPB medium were similar between the three strains (Fig. 9C). Following a phase of relatively fast growth during the first 2 days, growth slowed and cells entered the stationary phase after 7 days of growth (Fig. 9C). Notably, after 7 days of growth fibrils were observed in the supernatant of all cultures, which were remarkably long as compared to the size of the L-forms themselves. Interestingly, fibers in all strains could be stained with CFW (Fig. 9B), indicating that they contain β-(1,3)- or β-(1,4)-glycosidic bonds.

To determine the monosaccharide composition of the glycan(s) produced by the L-form strains, insoluble substances were isolated from the parental strain and the *cslA*$_{SV}$ and *glxA*$_{SV}$ deletion mutants in triplicate. Typically, from a 100 ml 7-day-old culture approximately 1 mg of product was obtained. However, for one wild-type sample and two *cslA*$_{SV}$ mutant samples the amounts were insufficient for analysis. The monosaccharide composition of the other samples is shown in Fig. 10. Notably, the fraction of GlcNAc in the *S. viridifaciens* L-form

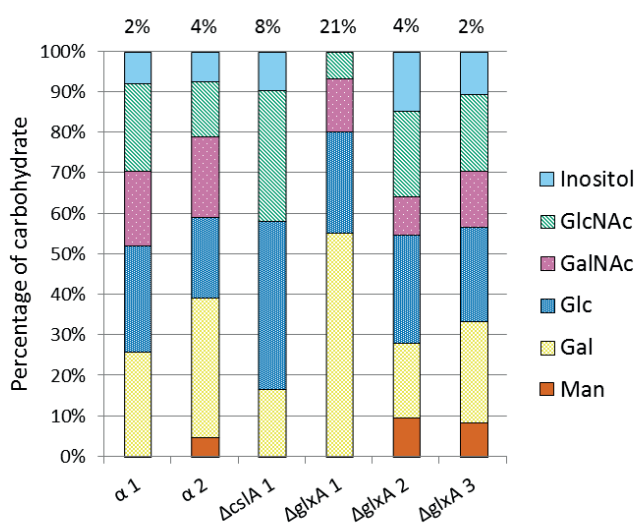

Figure 10. *Carbohydrate monomer composition of insoluble particles in L-forms cultures of S. viridifaciens alpha (parental strain), ΔcslA*$_{SV}$ *and ΔglxA*$_{SV}$ *as detected after acid hydrolysis and GLC-MS quantification of pentose and hexose sugars (colored as indicated in figure legend). Values are represented as a mole-based percentage of total recovered carbohydrate. Percentages above each bar represent the fraction of carbohydrate as a percentage of the total weight of the sample. The abbreviations used are: GlcNAc, N-acetylglucosamine; GalNAc, N-acetylgalactosamine; Glc, glucose; Gal, galactose; Man, mannose.*

samples varies between 10-30% compared to 60-90% for *S. lividans* growing as a mycelium (see Fig. 7). As a consequence, sugars that are derived from glycans which are produced in (very) low quantities compared to the abundance of peptidoglycan now make up a significant part of the total fraction. An example is galactose, which comprises 4 to 8% of the monosaccharides from the *S. lividans* mycelial samples, in contrast to the samples from L-forms where the percentage of galactose ranged from 16 to 55%. Excitingly, no GalNAc was detected in the sample obtained from the *cslA*$_{SV}$ mutant L-forms. This may indicate that the polymer produced by CslA contains GalNAc residues, an observation that awaits verification. The strong reduction in the levels of GlcNAc due to the absence of PG shows the potential of L-forms for the characterization of glycans produced in small amounts.

Discussion

Glycans are formed in all kingdoms of life, where they play pivotal roles during growth and development. They mostly occur in the extracellular space, where they contribute to providing structural integrity of the surrounded cells. Some polysaccharides, such as the cell wall constituent peptidoglycan, are abundantly synthesized, while others are produced at barely detectable levels. The characterization of glycan structures is often limited by the amounts and purity of the isolated glycan (Mulloy *et al.*, 2009). In this study we aimed to characterize a glycan produced by the *Streptomyces* CslA protein, which appears to represent only a minor fraction of the cell wall matrix. To this end, we developed a cell wall deficient L-form platform that due to the absence of PG simplifies glycan purification and characterization. Preliminary analyses using this innovative platform hints that the glycan produced by CslA may contain GalNac.

Limitations in bioinformatics and indirect approaches to determine glycosyltransferase specificity

The key enzymes in the production of glycans are the glycosyltransferases (GTs), which catalyze glycosidic linkage formation between the sugar moieties of activated nucleotide sugars and glycosyl acceptor molecules. Combining sequence- and network- based computational predictions can unveil insight in the donor and acceptor molecules, which can be growing glycan chains, lipids, peptidoglycan or proteins (Sánchez-Rodríguez *et al.*, 2014). Bioinformatics analysis shows that CslA$_{SL}$ is a processive glycosyltransferase that couples monosaccharides to a growing glycan chain, which is transported through the membrane into the extracellular environment. Genome synteny analysis supports the role of CslA$_{SL}$ as a polysaccharide producing protein, based on the observation that the *cslA*$_{SL}$ gene is surrounded by a number of genes that encode proteins functioning in carbohydrate metabolism. However, predicting the specificity of GTs is still complicated especially for GTs

of Gram-positive bacteria, for which clear sequence motifs determining substrate specificity are largely unknown (Weerapana and Imperiali, 2006). For example, the active site of $CslA_{SL}$ is very similar to the established bacterial cellulose synthases BcsA (Morgan *et al.*, 2013), but contains a number of amino acids substitutions in crucial motifs, which makes it unreliable to conclude that $CslA_{SL}$ synthesizes cellulose.

In addition to computational analyses, polysaccharides are often studied using indirect methods, including the use of fluorescent stains, antibodies, lectins and hydrolytic enzymes (Cummings and Etzler, 2009; Mulloy *et al.*, 2009). Likewise, $CslA_{SL}$ was predicted to produce cellulose, based on 1) the $CslA_{SL}$-dependent hyphal tip staining with calcofluor white (CFW), which stains β-(1,3)- and β-(1,4)-glycans including cellulose and chitin (Wood, 1980), and 2) the abolishment in hyphal attachment of the wild-type strain when grown in the presence of cellulase, which mimics the phenotype of a *cslA*$_{SL}$ mutant (Xu *et al.*, 2008; de Jong, Wösten, *et al.*, 2009). Our work shows that not only treatment with cellulases, but also treatment with chitinases is able to abolish attachment of hyphae to hydrophobic surfaces. These data can be interpreted in different ways: 1) the polymer produced by CslA may contain stretches of both cellulose and chitin; 2) the hydrolytic enzyme preparations are not pure or specific enough and may therefore hydrolyse different types of carbohydrates (Schiavone *et al.*, 2014); or 3) cellulose and chitin are produced during adhesion, both of which are required for attachment. The only way to discriminate between these options is to purify the glycan followed by characterization of its composition and structure.

L-forms as a platform to overcome the PG barrier for glycan characterization

Synthesis of the $CslA_{SL}$-dependent glycan occurs at growing hyphal tips (Xu *et al.*, 2008; de Jong, Wösten, *et al.*, 2009; Chaplin *et al.*, 2015), which are subject to constant remodeling during the process of tip extension, and which poses a threat to the integrity of the entire hypha. During tip growth, nascent layers of peptidoglycan are deposited outside the cell, which subsequently become cross-linked by transpeptidase activity, thereby leading to the formation of mature PG. The glycan produced by CslA is thought to provide protection by forming an extracellular bandage-like structure, which may protect the site where PG is still present in a nascent form (Chater *et al.*, 2010). Having fulfilled its protective role, the polymer may be degraded again sub-apically, for instance by the CslZ protein being part of the *cslA-glxA* gene cluster (Liman *et al.*, 2013). Cell-wall analysis with FTIR and NMR did not result in significant differences between *S. lividans* 1326 and its Δ*cslA* mutant. Furthermore, the monosaccharide analysis of the cell walls showed predominantly the sugar GlcNAc. These results strongly imply the high relative abundance of peptidoglycan, and are supportive for a model in which the $CslA_{SL}$-dependent glycan is only transiently present, or only comprises a marginal fraction of the total amount of glycan present at the cell surface. An L-form cell line that is derived from the strain of interest has the advantage

that precursors for glycan synthesis are created by the organism itself; other unknown essential proteins for glycan synthesis are conserved; and deletion mutants of homologues genes can be created for comparison. As expected, monosaccharide analysis revealed that the amount of GlcNAc in L-forms is dramatically reduced, indicating that the majority of PG is absent. The loss of the cell wall has an additional benefit as it leads to the secretion of otherwise cell wall-associated glycans in the culture broth, which greatly facilitates their purification and subsequent characterization. We therefore anticipate that this platform will be useful for studying GTs from many different organisms, in particular if the glycan only comprises a marginal fraction of the total amount of glycans produced by the natural host. In this context it is interesting to mention that the molecular techniques for manipulating L-forms have already been developed.

What is the composition of the CslA-dependent glycan?

The monosaccharide composition analysis of cell walls from WT and $cslA_{SL}$ and $glxA_{SV}$ deletion mutants suggests with a decrease in the amount of glucose upon gene deletion that CslA might indeed produce cellulose. In contrast, analysis of the L-form strains only showed the remarkable absence of GalNAc in the $cslA_{SV}$ mutant, while glucose remained constant. We could speculate that CslA produces a glucose based polymer with incorporation of other monosaccharides like GalNAc under certain conditions. This type of substrate promiscuity of GTs is not uncommon in bacteria as is demonstrated by the incorporation of GlcNAc by the cellulose synthase of *Glucanoactebacter xylinus* (Lee *et al.*, 2001; Sánchez-Rodríguez *et al.*, 2014). Interestingly, the active site of processive glycosyltransferases can give rise to two separate types of glycosidic bonds (May *et al.*, 2012), suggesting that CslA may produce a glycan with alternating glycosidic linkages. Overexpression of the the *cslA-glxA* gene cluster in the L-form platform would be the next step towards characterizing the CslA-produced glycan.

Materials and Methods

Bacterial strains and growth conditions

The bacterial strains used in this study are presented in Table 1. *E. coli* K12 strains JM109 (Sambrook *et al.*, 1989) and ET12567 (MacNeil *et al.*, 1992) were used for plasmid propagation and were grown and transformed using standard procedures (Sambrook *et al.,* 1989). Soy flour mannitol (MS; Kieser *et al.*, 2000) agar plates were used for the isolation of spores from *Streptomyces lividans* wild-type (1326) and mutant strains. For phenotypical characterization and growth of *Streptomyces viridifaciens* wild-type (DSM40239) and mutant strains, MYM agar plates (Jakeman *et al.*, 2006) or TSBS liquid medium were used.

30 ml liquid cultures were inoculated with 10^6 WT or Δ*cslA*$_{SV}$ spores per ml or with 1ml of precultured Δ*glxA*$_{SV}$ culture and analysed after 24h of growth.

For growth of all *S. viridifaciens* L-form strains, liquid L-phase broth (LPB) and solid L-phase medium (LPM) were used. LPB is a mixture of 50 % TSBS and 50 % YEME supplemented with 25 mM $MgCl_2$ (Kieser *et al.*, 2000). L-phase medium (LPM) contains 200 g sucrose, 5 g glucose, 5 g yeast extract, 5 g peptone, 0.1 g MgSO4, and 7.5 g Iberian agar per liter, and is supplemented with 25 mM $MgCl_2$ and 5 % horse serum (Sigma Aldrich) after autoclaving. L-form cultures were shaken at 100 rpm, and, if necessary, inoculated in fresh LPB (1:100 dilution) after one week.

Table 1 *Strains used in this study*

Strains	Description	Reference or source
***S. lividans* strains**		
1326	*Streptomyces lividans* 1326	(Hopwood *et al.,* 1985)
1326 Δ*cslA*$_{SL}$	*S. lividans* 1326 Δ*cslA*$_{SL}$ marker-less	(Chaplin *et al.*, 2015)
1326 Δ*glxA*SL	*S. lividans* 1326 Δ*glxA*$_{SL}$ marker-less	(Chaplin *et al.*, 2015)
***S. viridifaciens* mycelial strains**		
DSM40239	*Streptomyces viridifaciens* DSM40239	(Soliveri *et al.*, 1993)
DSM40239 Δ*cslA*$_{SV}$	*S. viridifaciens* DSM40239 *cslA*$_{SV}$*::ApraLoxP*	This study
DSM40239 Δ*glxA*$_{SV}$	*S. viridifaciens* DSM40239 *glxA*$_{SV}$*::ApraLoxP*	This study
***S. viridifaciens* L-form strains**		
α	*S. viridifaciens* DSM40239 L-form cell line	K. Ramijan, *unpublished*
α Δ*cslA*$_{SV}$	*S. viridifaciens* DSM40239 α *cslA*$_{SV}$*::ApraLoxP*	This study
α Δ*glxA*$_{SV}$	*S. viridifaciens* DSM40239 α *glxA*$_{SV}$*::ApraLoxP*	This study
***E. coli* strains**		
JM109	See reference	(Sambrook *et al.,* 1989)
ET12567	See reference	(MacNeil *et al.*, 1992)

Attachment assay

Attachment of hyphae to polystere surface was assessed and quantified as earlier described (de Jong, *et al.*, 2009). Briefly, 25-well plates were filled with 3 ml mNMMP (van Keulen *et al.*, 2003) containing 10^6 spores per ml. After the addition of spores, different amounts of chitinase from *Trichoderma viride* (Sigma Aldrich) or cellulase from *Trichoderma reesei* (Sigma Aldrich) were added. Plates were then sealed with parafilm to prevent dehydration. After 7 days of growth 100 µl of crystal violet solution (0.5%; Acros Organics) was added to each well and plates were left at room temperature for 10 min. Subsequently, the supernatant with free floating biomass was removed and plates were vigorously washed with running tap water and dried at 50 °C. Crystal violet was extracted from the attached biomass with 4

ml of 10% SDS during 30 min incubation at room temperature. The OD_{570} of 100 µl aliquots was determined using a spectrophotometer. For each enzyme treatment eight biological replicates were analysed and for the attachment of the *S. lividans* wild-type $cslA_{SL}$ and $glxA_{SL}$ mutants thirteen biological replicates were analysed. The relative attachment was calculated by normalizing the data with the mean attachment of the untreated or wild-type samples in the same 25-well plate.

Generation of S. viridifaciens ΔcslA$_{SV}$ and ΔglxA mutants

The $cslA_{SV}$ and $glxA_{SV}$ null-mutants were created in *S. viridifaciens* DSM40239, and in the stable L-form derivative strain called α. The $cslA_{SV}$ and $glxA_{SV}$ genes were replaced via homologous recombination, with an apramycin resistance cassette (*aacC4*) flanked with *loxP* sequences using the unstable pWHM3 plasmid (Świątek *et al.*, 2012). In the $cslA_{SV}$ null-mutants nucleotides +195 to +2037 relative to the start of $cslA_{SV}$ were replaced, while in the $glxA_{SV}$ mutants nucleotides +164 to +1587 relative to the start of $glxA_{SV}$ were replaced. L-forms were transformed using an adjusted protocol for protoplast transformation (Kieser *et al.*, 2000): L-forms from 1 ml of a three days-old LPM culture were washed with, and subsequently resuspended in 1 ml P-buffer. For each transformation 50 µl of washed L-forms were mixed with 0.5 µg of plasmid DNA and 200 µl 25% PEG in P-buffer, after which the suspension was plated on LPM agar plates. Following transformation, mutants were selected that were thiostrepton-sensitive and apramycin-resistant. All mutants were verified by PCR analysis (see Table 2 for primer sequences).

Growth of L-forms

Growth of L-form cultures was measured using a spectrophotometer (BioRad) operated at a wavelength of 600 nm. Measurements were corrected for the absorbance of a (unknown) green pigment that is produced by L-forms, by subtracting the OD_{600} measurement of the supernatant obtained after 10 min centrifugation at 16,000 *g*.

Microscopy of L-forms

Morphology of L-forms was analysed by light microscopy using an Axioplan 2 (Zeiss) equipped with a DKC-5000 digital camera (Sony). Insoluble fibers were stained by incubating 9 µl of L-form culture with 1 µl calcofluor white solution (Sigma Aldrich) for 5 min. Stained samples were analysed using fluorescence microscopy equipped with a mercury lamp and Omega Optical Filter set XF06.

Table 2 *Primers used in this study*

Primer name	Primer sequence	Restriction site
CslA-P1-HindIII	GACAAGCTTAAGGAGGTGAGGGAGTTC	*Hind*III
CslA-P2-XbaI	GACTCTAGAGAAGTGACCGTAGTCATAGG	*Xba*I
CslA-P3-XbaI	GACTCTAGAACCACGACGAACAGCGACAC	*Xba*I
CslA-P4-EcoRI	GACGAATTCTCGGGTCGTAGACCTCGTTG	*Eco*RI
CslA-INT-FW	GGAACCGCAAGCACAACGTGAG	
CslA-INT-RV	AACACCACCGCCCAGAACAG	
Delcheck CslA-FW	GCCGTCCCAACCGTACAAG	
Delcheck CslA-RV	CACGTACTGGTGGTACTG	
Apra-primer	ATTCCGGGGATCCGTCGACC	
GlxA-P1-FW	GACGAATTCAGAATCCGCTCGCTCCAGTC	*Eco*RI
GlxA-P2-REV	GACTCTAGAAGTGGCCGTACTGTCCCTTG	*Xba*I
GlxA-P3-FW	GACTCTAGAGAGCAGCGCATCGAGATCTAC	*Xba*I
GlxA-P4-REV	GATAAGCTTAGGTCGGCTTCTCCTACCAG	*Hind*III
GlxA-INT-FW	ATGTTCTGCGGCGGTCACAC	
GlxA-INT-REV	TGCCCTGGTAGTCCTGCTTG	
Delcheck GlxA-FW	ACGACGAACAGCGACAC	
Delcheck GlxA-RV	CGTCCGTCAGGAACAACATGTACC	

Bioinformatics analyses

Genetic synteny analysis was performed with SyntTax (Oberto, 2013), using the CslA sequence of *S. lividans* ($CslA_{SL}$, accession number: EOY47900) as input. Protein topology predictions were made with TMHMM version 2.0 (Krogh *et al.*, 2001). Protein sequence comparisons were made by aligning $CslA_{SL}$ with BcsA of *Rhodobacter spaeroides* (WP_041669585.1), CelA of *Agrobacterium tumefaciens* (NP_533806.1) and $CslA_{AX}$ of *Acetobacter xylinus* (also known as *Komagataeibacter xylinus*, CAA38487.1) with Clustal Omega (Sievers *et al.*, 2011). The 3D-structure of $CslA_{SL}$ was predicted using Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2; Kelley *et al.*, 2015). The crystal structure of BcsA (c4hg6A) formed the highest scoring template and was used for homology modeling of $CslA_{SL}$ omitting the low confidence regions with high disorder prediction, which are the N-terminus (amino acid 1-78) and C-terminus (amino acid 590-653). The 3D-model was checked for accuracy with the build in Phyre investigator, showing that the general fold of the 3D-model of $CslA_{SL}$ is reasonable.

FTIR spectroscopy

Fourier Transform Infrared (FTIR) spectroscopy was performed on cell walls of the *S. lividans* wild-type strain and the Δ$cslA_{SL}$ mutant. Therefore, total mycelium obtained from 5 ml

broth of 24 h TSBS cultures was washed twice with demineralized water and sonicated (5s on/5s off) until the mycelium was broken (as visualized using microscopy). The insoluble cell walls were obtained by centrifugation, and subsequently washed with demineralized water, acetone and ether before being lyophilized (Lee *et al.*, 2001). 2 mg of cell walls were manually blended with 100 mg of dried KBr powder and pressed into a tablet. FTIR transmittance spectra of both samples were recorded with a Bio-Rad (Excalibur series) FTS4000 spectrometer between 4,000-500 cm^{-1} (64 scans, resolution 2 cm^{-1}) and corrected for background signals.

Solid state CPMAS-NMR

For NMR analysis mycelium was harvested from 24h TSBS grown cultures and lysed with a French pressure cell (Stansted Fluid Power Ltd.). Cell walls were harvested by centrifugation at 16,000 *g*, washed three times with 10 mM Tris-HCL pH 7.0 and freeze-dried overnight. Solid-state ^{13}C NMR studies were performed on a Bruker AV-750 spectrometer with a 17.6 Tesla magnetic field, in which carbon nuclei resonate at 188.64 MHz, respectively. A triple resonance MAS probe head of 4 mm with a standard ZrO_2 rotor was used to spin till 13 kHz. For ^{1}H to ^{13}C CPMAS with TPPM decoupling, we used 49.1 kHz and a ramp from 80 kHz to 100kHz RF field frequencies in the ^{13}C and ^{1}H channels, respectively, for CP. The contact time was 2 ms, and the repetition time was 1 s for 512 scans. Chemical shift references (0 ppm) were externally referenced with TMS for ^{1}H and ^{13}C.

Monosaccharide analysis by GLC-MS

For monosaccharide analysis, mycelium from 30 ml 24h TSBS grown cultures was collected and lysed with a French pressure cell (Stansted Fluid Power Ltd.). Cell walls were harvested by centrifugation at 12,000 *g* for 30 min and resuspended in DNAse buffer (10 mM Tris-HCl, 2.5mM $MgCl_2$, 0.5mM $CaCl_2$, pH7.6). Cell walls were treated with 5 µl DNAse and 5 µl RNAse (Thermo Scientific) for 1 h at 30 ˚C. After 20 min centrifugation at 12,000 *g,* pellets were resuspended in protein extraction buffer (2% SDS, 40mM β-mercaptoethanol, 50 mM Tris/HCl, 5 mM EDTA, pH7.4; Gastebois *et al.*, 2010; van Munster *et al.*, 2013) and incubated for 10 min at RT. After centrifugation for 20 min at 12,000 *g,* cell walls were washed twice with demineralized water and lyophilized.

Fibers produced by L-forms were isolated by spinning down (12,000 *g* for 30 min) insoluble material present in 50 ml of 7-day-old liquid-grown cultures. The pellet fraction, containing L-forms and fibers, was resuspended in DNAse buffer, thereby lysing the L-forms. The lysate was further processed as described above. Samples containing a relatively high level of deoxyribose indicative of ineffective DNAse and RNAse treatment were excluded from

further analysis.

The method for monosaccharide analysis was previously described by Van Leeuwen *et al.* (2008). The polysaccharide containing samples were subjected to methanolysis (1.0 M methanolic HCl, 24 h, 85 °C), followed by re-*N*-acetylation and trimethylsilylation (1:1:5 hexamethyldisilazane–trimethylchlorosilane–pyridine; 30 min, room temperature). The mixture of trimethylsilylated (methyl ester) methyl glycosides were analysed by GLC on an ZB-1HT column (30 m × 0.25 mm; Phenomenex, Utrecht, The Netherlands), using a Trace 1300 gas chromatograph with flame-ionization detection (Fisher Scientific, Amsterdam, The Netherlands; temperature program 140 – 240 °C, 4 °C/min). The identification of the monosaccharide derivatives was confirmed by GLC-EI-MS analysis on a QP2010 Plus system (Shimadzu, 's-Hertogenbosch, The Netherlands), using the same column and temperature program.