New insights into Nod factor biosynthesis: Analyses of chitooligomers and lipo-chitooligomers of *Rhizobium* sp. IRBG74 mutants

Véréna Poinsot [a, 1], Matthew B. Crook [b, 1], Stéphanie Erdn [c], Fabienne Maillet [c], Adeline Bascaules [a], Jean-Michel Ané [b, d, *]

[a] *Laboratoire des IMRCP, UMR 5623 Université Paul Sabatier, Toulouse, France*
[b] *Department of Bacteriology, University of Wisconsin–Madison, Madison, WI 53706, USA*
[c] *Laboratoire des Interactions Plantes–Microorganismes, UMR 441/2594 INRA-CNRS, B.P. 52627, F-31326 Castanet-Tolosan Cedex, France*
[d] *Department of Agronomy, University of Wisconsin–Madison, Madison, WI 53706, USA*

ARTICLE INFO

ABSTRACT

Soil-dwelling, nitrogen-fixing rhizobia signal their presence to legume hosts by secreting lipo-chitooligomers (LCOs) that are decorated with a variety of chemical substituents. It has long been assumed, but never empirically shown, that the LCO backbone is synthesized first by NodC, NodB, and NodA, followed by addition of one or more substituents by other Nod proteins. By analyzing a collection of in-frame deletion mutants of key *nod* genes in the bacterium *Rhizobium* sp. IRBG74 by mass spectrometry, we were able to shed light on the possible substitution order of LCO decorations, and we discovered that the prevailing view is probably erroneous. We found that most substituents could be transferred to a short chitin backbone prior to acylation by NodA, which is probably one of the last steps in LCO biosynthesis. The existence of substituted, short chitin oligomers offers new insights into symbiotic plant–microbe signaling.

1. Introduction

Rhizobia are soil-dwelling, Gram-negative bacteria that can establish a nitrogen-fixing symbiosis with legumes [1]. In this association, legumes develop specialized organs, root nodules, to host the bacteria under low oxygen tensions that are favorable to nitrogen fixation and provide hydrocarbons (malate) to the microorganism to sustain, in particular, the nitrogenase activity. In return, the rhizobia provide ammonium and amino-acids, transformed from atmospheric dinitrogen, to the plant host.

During the early stages of this symbiosis, a molecular dialogue takes place. The plant exudes flavonoids and isoflavonoids into the rhizosphere, and the rhizobia respond by secreting Nod factors [2]. Nod factors are lipo-chitooligomers (LCOs) that are required for

nodule development and intracellular infection of the root and are active at very low (nanomolar to picomolar) concentrations. Nod factor structures were extensively studied in the 1990s, and their biosynthesis has also been investigated based on molecular biological tools [3,4]. The enzymes responsible for Nod factor biosynthesis are encoded by *nod*, *noe*, and *nol* genes, which are located either on symbiotic plasmids or in genomic islands. These genes are found under the control of regulatory elements, called *nod* boxes [5], which are recognized by the transcription factor NodD [6].

While forward and reverse genetics have allowed us to ascertain the role of all the genes involved in Nod factor biosynthesis [7], little is known about the acting order of their corresponding enzymes. To explore this question, we analyzed two classes of molecules, LCOs and chitooligomers (COs), produced by wild-type, overproducing, and mutant strains of *Rhizobium* (*Agrobacterium*) sp. IRBG74 (hereafter IRBG74). This strain was chosen because the annotation of its genome [8] suggested that it would produce a wide array of LCOs, in contrast to more commonly studied strains, such as *Sinorhizobium meliloti* 1021, which produce limited numbers of LCO structures [3]. This strain is also of agronomic interest because it has been shown to colonize rice and increase plant
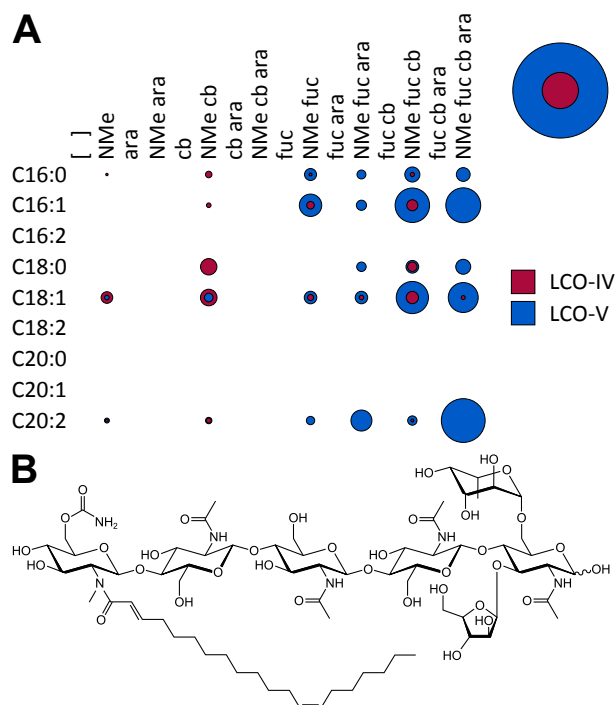
A

B

Fig. 1. A. A bubble matrix of LCOs produced by wild-type *Rhizobium* sp. IRBG74. ara, arabinose; cb, carbamoyl; fuc, fucose; NMe, *N*-methyl. Relative abundances of different chemical species are represented by the size of the bubble where a given row and column overlap. The bubbles at the top right of the figure indicate the overall abundances of the indicated classes of chemicals using the same scale. Fig. 2 and 5–11 follow this same format. **B**. The structure of the major LCO produced by wild-type *Rhizobium* sp. IRBG74.

growth and yields [9,10] in addition to nodulating *Sesbania* species [11], which are recommended for use as a green manure in southeast Asia. With the help of LC-MS/MS experiments, we have been able to deduce the potential step-by-step progression of the biosynthesis of the COs and LCOs of this strain.

2. Results and discussion

2.1. IRBG74 produces a wide family of LCOs and COs

2.1.1. The LCOs of IRBG74

Wild-type IRBG74 does not secrete LCOs in substantial amounts. The analysis of 5 L of bacterial exudates was not sufficient for the detection of LCOs by MS (data not shown). Because LCOs are amphipathic, they are most likely to be found in the lipid bilayer [12]. Thus to enhance the production of LCOs, plasmid pA28 [13], which overexpresses the regulatory $nodD_1$ gene of *Sinorhizobium* sp. NGR234, was introduced into IRGB74. The resulting strain secreted amounts of LCOs that were sufficient for analysis. The LCOs secreted into the culture medium by overexpressing strains are not expected to differ significantly from those found in the membranes of wild-type rhizobia [14]. More than 40 different structures were found in the culture supernatant of the *nodD*-overexpressing strain (Fig. 1A). All LCOs exhibited N-methylation of the non-reducing end. Most species were five GlcNAc residues long, carbamoylated, and fucosylated, while arabinosylation was more variable (Fig. 1B). Unsaturated acyl chains were more prevalent, with C16:1 and C18:1 occurring in nearly equal amounts (Table 1). Notably, arabinosylation never occurred in the absence of fucosylation, and more than half of the LCO-IV molecules were carbamoylated without any other decoration.

To check whether the composition of IRBG74's LCOs was affected by regulation using the *nodD* of a phylogenetically distant strain, the native IRBG74 regulator $nodD_1$ and the $nodD_2$–*syrM* regulatory module were each cloned into a plasmid under the control of a constitutive promoter and introduced into wild-type IRBG74. We confirmed by HPLC that the *nodD* genes of *Rhizobium* sp. IRBG74 could be induced by apigenin (Fig. S1) The LCOs produced by overexpression of each of the native *nodD* genes of IRBG74

Table 1
Relative percentages of various features of COs and LCOs produced by *Rhizobium* sp. IRBG74 and various *nod* mutants.

|	WT LCOs	WT COs	*nodC* COs	*nodZ* LCOs	*nodB* COs	*nodS* LCOs	*nodU* LCOs	*nodA* COs	*nodE* LCOs	*noeP* LCOs	*noeN* LCOs	*p0345* LCOs
II	0.0%	7.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.5%	0.0%	0.0%	0.0%	0.0%
III	0.0%	34.3%	0.0%	0.0%	37.8%	9.9%	0.0%	2.7%	0.0%	0.0%	0.0%	0.0%
IV	12.6%	39.1%	0.0%	42.4%	48.6%	51.6%	9.8%	66.0%	24.2%	7.8%	19.1%	19.3%
V	87.4%	19.3%	0.0%	57.6%	13.6%	38.6%	90.2%	28.9%	75.8%	92.2%	80.9%	80.7%
VI	0.0%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
CO	0.0%	51.5%	0.0%	0.0%	100.0%	0.0%	0.0%	7.9%	0.0%	0.0%	0.0%	0.0%
dCO	0.0%	48.5%	0.0%	0.0%	0.0%	0.0%	0.0%	92.1%	0.0%	0.0%	0.0%	0.0%
C16:0	6.9%	0.0%	0.0%	0.0%	0.0%	6.9%	0.0%	0.0%	10.4%	0.8%	11.6%	0.8%
C16:1	31.0%	0.0%	0.0%	19.8%	0.0%	20.5%	12.7%	0.0%	41.8%	10.5%	23.6%	29.6%
C16:2	0.0%	0.0%	0.0%	13.3%	0.0%	1.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
C18:0	8.0%	0.0%	0.0%	29.0%	0.0%	16.1%	35.1%	0.0%	5.0%	27.0%	0.9%	12.4%
C18:1	28.7%	0.0%	0.0%	0.0%	0.0%	30.3%	28.3%	0.0%	42.8%	36.0%	42.7%	32.4%
C18:2	0.0%	0.0%	0.0%	0.0%	0.0%	2.9%	0.0%	0.0%	0.0%	0.0%	7.1%	0.0%
C20:0	0.0%	0.0%	0.0%	0.0%	0.0%	1.1%	0.0%	0.0%	0.0%	0.0%	3.0%	0.0%
C20:1	0.0%	0.0%	0.0%	18.6%	0.0%	8.5%	0.0%	0.0%	0.0%	22.9%	3.2%	13.9%
C20:2	25.3%	0.0%	0.0%	19.3%	0.0%	12.7%	23.9%	0.0%	0.0%	2.7%	7.9%	10.9%
no methyl	0.0%	54.5%	0.0%	0.0%	100.0%	100.0%	0.0%	18.4%	0.0%	0.0%	0.0%	0.0%
methyl	100.0%	45.5%	0.0%	100.0%	0.0%	0.0%	100.0%	81.6%	100.0%	100.0%	100.0%	100.0%
no fucose	9.1%	49.8%	0.0%	94.6%	54.4%	17.8%	0.0%	15.6%	17.0%	0.0%	6.7%	9.7%
fucose	90.9%	50.2%	0.0%	5.4%	45.6%	82.2%	100.0%	84.4%	83.0%	100.0%	93.3%	90.3%
no carbamoyl	19.9%	60.2%	0.0%	10.8%	100.0%	68.7%	100.0%	24.7%	6.7%	12.9%	20.4%	12.6%
carbamoyl	80.1%	39.8%	0.0%	89.2%	0.0%	31.3%	0.0%	75.3%	93.3%	87.1%	79.6%	87.4%
no arabinose	48.3%	100.0%	0.0%	99.1%	100.0%	58.7%	57.1%	83.2%	75.3%	100.0%	72.2%	83.8%
arabinose	51.7%	0.0%	0.0%	0.9%	0.0%	41.3%	42.9%	16.8%	24.7%	0.0%	27.8%	16.2%
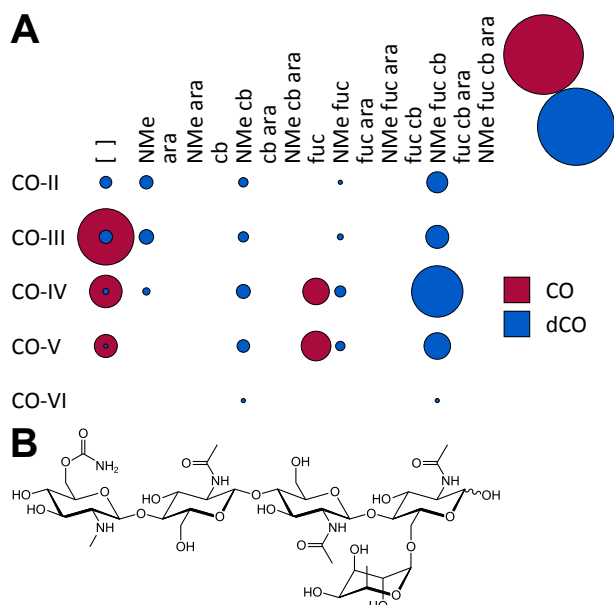
Fig. 2. **A**. A bubble matrix of COs and dCOs produced by wild-type *Rhizobium* sp. IRBG74. Abbreviations as in Fig. 1. **B**. The structure of the major dCO produced by wild-type *Rhizobium* sp. IRBG74.

Fig. 4. Relative production of COs, dCOs, and LCOs by WT, *nodA*, *nodB*, *nodC*, and *nodABC* strains of *Rhizobium* sp. IRBG74. Values are given in counts per scan (intensity of the detection) assuming that the concentration of each extract was the same (0.1 g/L).

were very similar to each other but were noticeably different from the profile of the strain overexpressing the *nodD₁* of *Sinorhizobium* sp. NGR234 (Fig. S2). Because the latter produced a greater diversity of LCO structures, subsequent analyses were conducted exclusively using this overexpression plasmid.

2.1.2. The short COs of IRBG74

For all the overproducing strains, the bacterial growth medium also contained short COs. Historically these molecules have been discarded during the organic extraction of LCOs, making this the first study to look at them in depth. The COs of IRBG74 ranged in degree of polymerization from two to five GlcNAc residues long, with CO-III and CO-IV being the most common (Fig. 2A). Roughly half of the COs detected in the wild-type exudates were mono-deacetylated at the non-reducing end (Table 1). These were almost always *N*-methylated, carbamoylated, and fucosylated but never arabinosylated (Fig. 2B).

2.2. Analysis of nod mutants

2.2.1. Overview

To investigate the order in which the various *nod* genes contribute to the final LCO structures, we developed a collection of mutants in IRBG74. The recent publication of the IRBG74 genome [8] enabled us to identify two clusters of *nod* genes on the symbiosis plasmid, pIRBG74a, and assign putative functions to many of the annotated *nod* genes by homology with previously described *nod* genes (Fig. 3). To ensure that there were no polar effects on
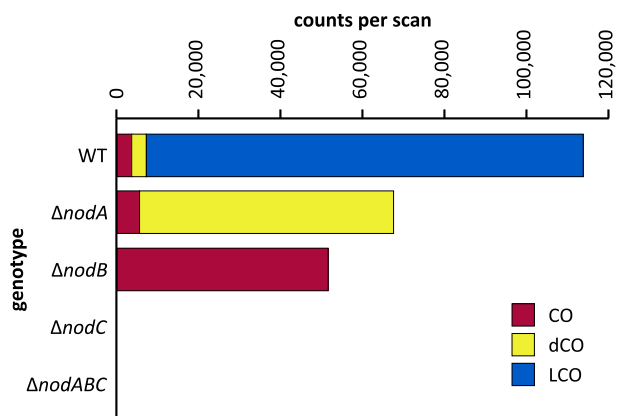
expression of downstream genes, we constructed an in-frame deletion of each target gene. In cases where multiple *nod* genes are involved in the biosynthesis of a chemical precursor, we only deleted the transferase responsible for adding the precursor in order to avoid redundancy (*e.g.* we only deleted the arabinosyl-transferase *noeP* instead of each gene of the *noeCNHOP* arabinosylation operon individually). We started our analysis by deleting the "common" *nod* genes: *nodA*, *nodB*, and *nodC*. These three genes are found in all rhizobia examined so far, with the exception of the photosynthetic bradyrhizobia, which nodulate their *Aeschynomene* hosts without producing Nod factors [15]. These three genes are usually, but not always, found in a single operon, *nodABC*, which has engendered the prevailing model that the products of these genes act first, followed by elaboration with various chemical adducts from the various host-specific *nod* genes.

2.2.2. Chitin synthesis

NodC is an *N*-acetylglucosaminyltransferase (EC 2.4.1.–), located in the inner membrane [16], that catalyzes the formation of β1→4 glycosidic linkages between GlcNAc residues from the non-reducing end of the growing chitin chain [17,18]. It is also involved in determining the degree of polymerization of COs and LCOs [19]. The *nodC* and *nodABC* mutants of IRBG74 do not produce detectable amounts of COs or LCOs (Fig. 4), indicating that biosynthesis of these molecules has been abolished in these strains. To corroborate this result we also tested a *nodC* mutant of *Sinorhizobium meliloti*; again no COs or LCOs were detectable by MS (data not shown).

2.2.3. Mono-deacetylation

NodB is a chitin *N*-deacetylase (EC 3.5.1.–) which removes the acetyl group of only the non-reducing terminal GlcNAc [20]. As expected, the *nodB* mutant of IRBG74 still produces COs, but no longer produces dCOs or LCOs (Figs. 4 and 5). Surprisingly, substantial quantities (~46%) of CO-IV and CO-V with a fucose on the
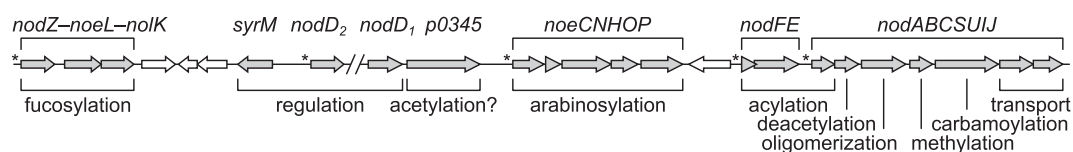

Fig. 3. The *nod* genes (depicted in gray) of *Rhizobium* sp. IRBG74 are organized into two clusters, both located on the symbiosis plasmid, pIRBG74a. Above the diagram gene or operon names are indicated. Below the diagram functions are indicated. Asterisks indicate the location of *nod* boxes.
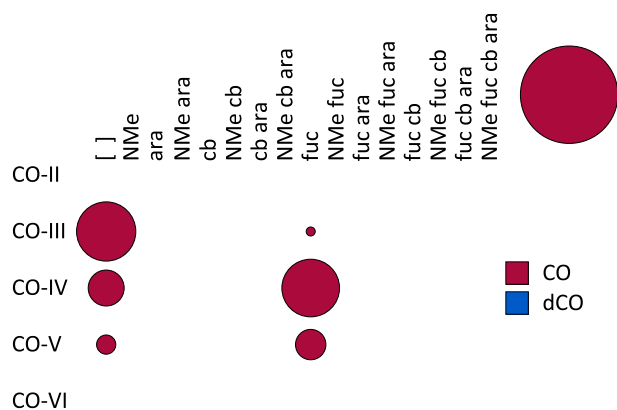
Fig. 5. A bubble matrix of COs produced by a *nodB* mutant of *Rhizobium* sp. IRBG74. Abbreviations as in Fig. 1.

non-reducing terminal GlcNAc were detected. This composition is similar to the COs produced by wild type (~50% fucosylated), but less than the LCOs produced by wild type (~91%) (Table 1). Methylation, carbamoylation, and arabinosylation were never detected. The degree of polymerization of COs was also reduced in the *nodB* mutant, with higher proportions of CO-III and CO-IV than what was observed in the wild type.

2.2.4. N-acylation

NodA is an *N*-acyltransferase (EC 2.3.1.—) that transfers fatty acids from an acyl carrier protein [21,22] to the non-reducing terminal GlcNAc of LCOs [23—25]. The *nodA* mutant of IRBG74 no longer produces LCOs, but it does produce a small distribution of short COs and a wide family of substituted dCOs, principally dCO-IV (Cb,Fuc,NMe) and dCO-V (Cb,Fuc,NMe) (Fig. 4 and Fig. 6A). Consistent with the results from the *nodB* mutant, COs are never *N*-methylated, arabinosylated, or carbamoylated unless they have been deacetylated first, but they can still be fucosylated. The dCOs of the *nodA* mutant are less likely to be *N*-methylated (~82% vs. 100%) or arabinosylated (~17% vs. ~52%) than the LCOs of wild-type (Table 1). However, compared to the dCOs produced by wild-type IRBG74, the rate of substitution increased for every type of substitution in the *nodA* mutant, except arabinosylation (Table 1). To corroborate these results, we also tested a *nodA* mutant of *Sinorhizobium meliloti*. Likewise, this strain no longer produced any LCOs but still produced both COs and dCOs (Fig. S3). Intriguingly, the COs showed a small degree of sulfation (~5%), which considerably increased in the dCO population (~63%). This is consistent with previous reports that the sulfotransferase, NodH, utilizes both COs and LCOs as substrates, though LCOs are preferred [26—28]. It is also worth noting that acetylation of the non-reducing terminal GlcNAc was never observed in the absence of sulfation (Fig. S3), suggesting that the former process may be dependent on the latter.

NodE is a specific acyl carrier protein that, along with NodF, a β-ketoacyl ACP synthase I (EC 2.3.1.41), is involved in the synthesis of polyunsaturated acyl groups which are then transferred to the non-reducing terminal GlcNAc of LCOs [29]. Deletion of *nodE* in IRBG74 predictably results in the complete loss of LCOs with polyunsaturated acyl groups and of LCOs with acyl chains longer than 18 carbons (Fig. 6B). There was also a mild decline in the proportion of arabinosylated LCOs (Table 1).

2.2.5. Fucosylation

Because we observed fucosylated COs produced by the *nodB* mutant of IRBG74, we next turned our attention to *nodZ*. NodZ is an α-1,6-fucosyltransferase (EC 2.4.1.—) that fucosylates C-6 of the
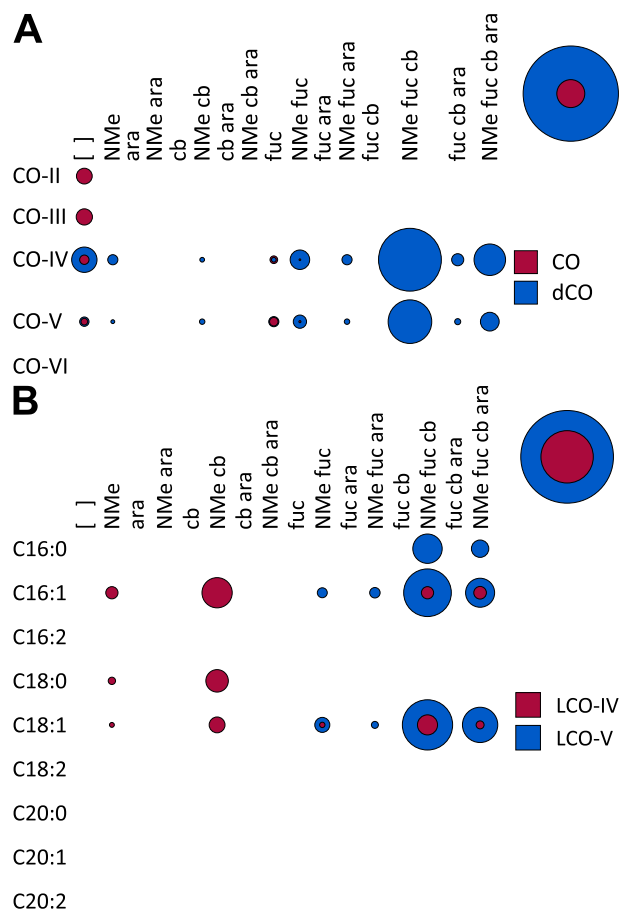

Fig. 6. A. A bubble matrix of COs and dCOs produced by a *nodA* mutant of *Rhizobium* sp. IRBG74. **B**. A bubble matrix of LCOs produced by a *nodE* mutant of *Rhizobium* sp. IRBG74. Abbreviations as in Fig. 1.

reducing terminal GlcNAc of Nod factors [30]. In the *nodZ* mutant fucosylation was not completely abolished—about 5% of LCOs were still fucosylated (Table 1). A similar result was observed for a *nodZ* mutant of *Azorhizobium caulinodans* ORS571 [31]. The persistence of fucosylation in the absence of NodZ may be due to the presence of an alternative, albeit less specific, fucosyltransferase in the IRBG74 genome or may indicate that the activation energy of fucose reacting with C-6 of the terminal GlcNAc of (L)COs is low enough that this reaction naturally occurs at a low level. No LCOs bearing arabinose groups directly on the chitin backbone could be detected unless there was also a fucose present (Fig. 7), which was also observed for a *nodZ* mutant of *Azorhizobium caulinodans* ORS571 [31]. N-methylation was still universal and carbamoylation still common in the *nodZ* mutant.

2.2.6. N-Methylation

NodS is a SAM-dependent methyltransferase (EC 2.1.1.—) that *N*-methylates the non-reducing terminal GlcNAc of Nod factors [32,33]. As expected, the LCOs of the *nodS* mutant were strictly unmethylated at the non-reducing terminal GlcNAc (Fig. 8). Surprisingly, the LCOs of the *nodS* mutant showed a low degree of carbamoylation. In wild-type IRBG74, 80% of the LCOs are carbamoylated whereas in the *nodS* mutant this was reduced to about 30% (Table 1). Arabinosylation and fucosylation were unaffected. There was also a reduction in the degree of polymerization compared to wild type (Table 1), similar to what others have observed [34].
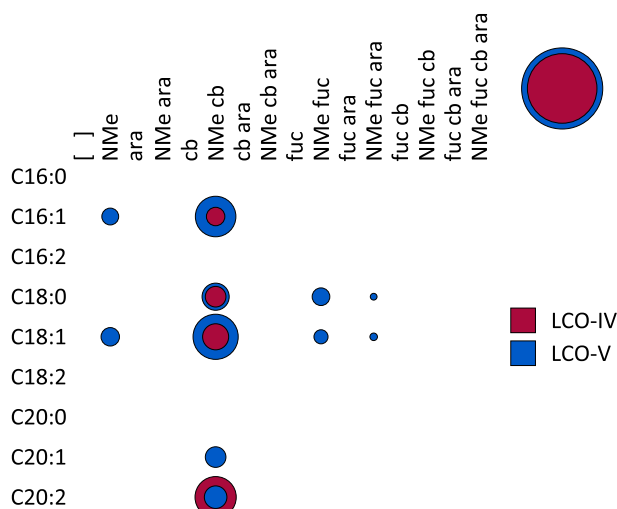
Fig. 7. A bubble matrix of LCOs produced by a *nodZ* mutant of *Rhizobium* sp. IRBG74. Abbreviations as in Fig. 1.
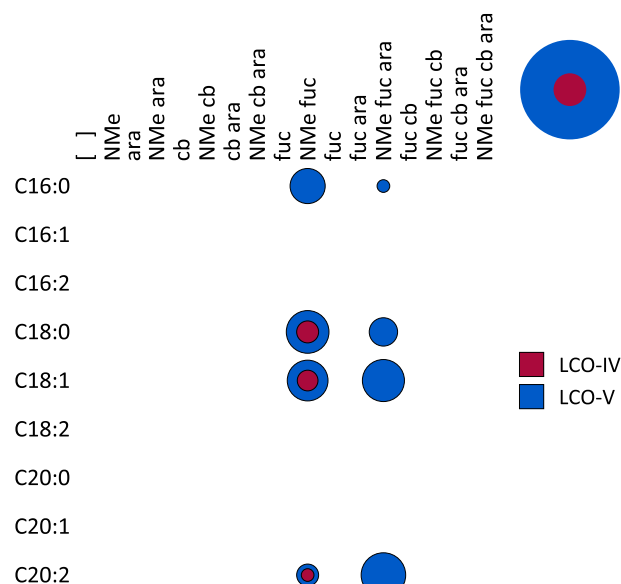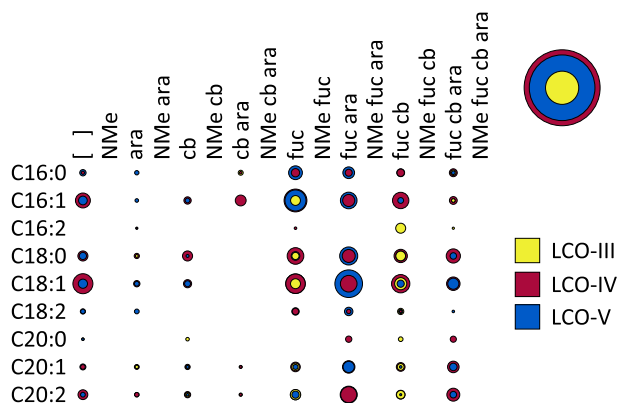

Fig. 8. A bubble matrix of LCOs produced by a *nodS* mutant of *Rhizobium* sp. IRBG74. Abbreviations as in Fig. 1.

2.2.7. Carbamoylation

NodU is an *O*-carbamoyltransferase (EC 2.1.3.−) that carbamoylates C-6 of the non-reducing terminal GlcNAc of Nod factors [34]. A second *O*-carbamoyltransferase found in some rhizobia (but not IRBG74), NolO, carbamoylates C-3 of the non-reducing terminal GlcNAc of Nod factors [35]. The LCO profile of the *nodU* mutant of IRBG74 was similar to wild type except that no carbamoylated species were observed (Table 1). Also, in the absence of carbamoylation, LCO-IV species were never arabinosylated (Fig. 9).

2.2.8. Arabinosylation

NoeP was first described as part of the *noeCHOP* operon of *Azorhizobium caulinodans* ORS571, which is involved in D-arabino-furanosylation of C-3 of the reducing-end GlcNAc of Nod factors [31]. Later work on orthologs in *Mycobacterium tuberculosis* established that *noeC*, *noeH*, and *noeO* were involved in three of the four steps in the biosynthesis of decaprenylphosphoryl arabinofuranose [36,37]. No orthologs of *noeP* were detected in *Mycobacterium tuberculosis* [36], suggesting it encodes the 3-*O*-arabinosyltransferase (EC 2.4.2.−). When *noeP* is deleted in IRBG74, no arabinosylated LCOs are produced (Fig. 10A). Fucosylation, carbamoylation, and methylation of LCOs are unaffected (Table 1).

The arabinosylation operon of IRBG74 includes an additional gene, *BN877_p0343*, which here we propose to name *noeN*. The

Fig. 9. A bubble matrix of LCOs produced by a *nodU* mutant of *Rhizobium* sp. IRBG74. Abbreviations as in Fig. 1.

noeN gene may encode the decaprenylphosphoryl-β-D-5-phosphoribose phosphatase (EC 3.1.3.−) not identified by Mikusová et al. [37] Deletion of *noeN* in IRBG74 reduces, but does not abolish, arabinosylation of LCOs (Fig. 10B), suggesting that the reaction catalyzed by NoeN (possibly conversion of *trans,octacis*-decaprenylphospho-β-D-ribofuranose 5-phosphate to *trans,octacis*-decaprenylphospho-β-D-ribofuranose) can occur spontaneously or else can be inefficiently performed by some other enzyme produced by IRBG74. As with the *noeP* mutant, fucosylation, carbamoylation, and methylation of LCOs are unaffected in the *noeN* mutant (Table 1).

2.2.9. BN877_p0345

Adjacent to the *nodD₁* gene of IRBG74 is a gene systematically annotated as *BN877_p0345* (hereafter simply *p0345*; Fig. 4) that encodes a putative acyltransferase or acetyltransferase (EC 2.3.−.−). It shares 28.5% identity with *noeT* of *Neorhizobium galegae* HAMBI 1141 and 28.1% identity with *hsnT* of *Rhizobium leguminosarum* sv. *trifolii* ICC105 at the amino acid level. NoeT is responsible for *O*-acetylation of C-3 of the GlcNAc residue immediately adjacent to the non-reducing terminal GlcNAc [38,39]. The *hsnT* gene is frequently found associated with *nodF* and *nodE*, suggesting that it may encode an acyltransferase [40,41]. In the *p0345* mutant we did not observe the complete or near complete eradication of any LCO modification (Fig. 11), so its function is still unknown. We did, however, observe a lower degree of arabinosylation in the *p0345* mutant (~16%) compared to the wild type (~52%) (Table 1).

3. Conclusion

It has generally been assumed (see, for example, Jabbouri et al.[34]), based on the fact that *nodA*, *nodB*, and *nodC* occur in a single operon in most rhizobia examined, that the basic LCO structure is synthesized first (chain elongation → deacetylation → acylation), and then other chemical substituents are added later (Scheme 1). Our results suggest that this view may be erroneous.

NodC necessarily acts first to polymerize UDP-GlcNAc into chitin, mainly CO-IV and CO-V (Fig. 4). However, after this point our results suggest that the synthesis diverges from the traditional view
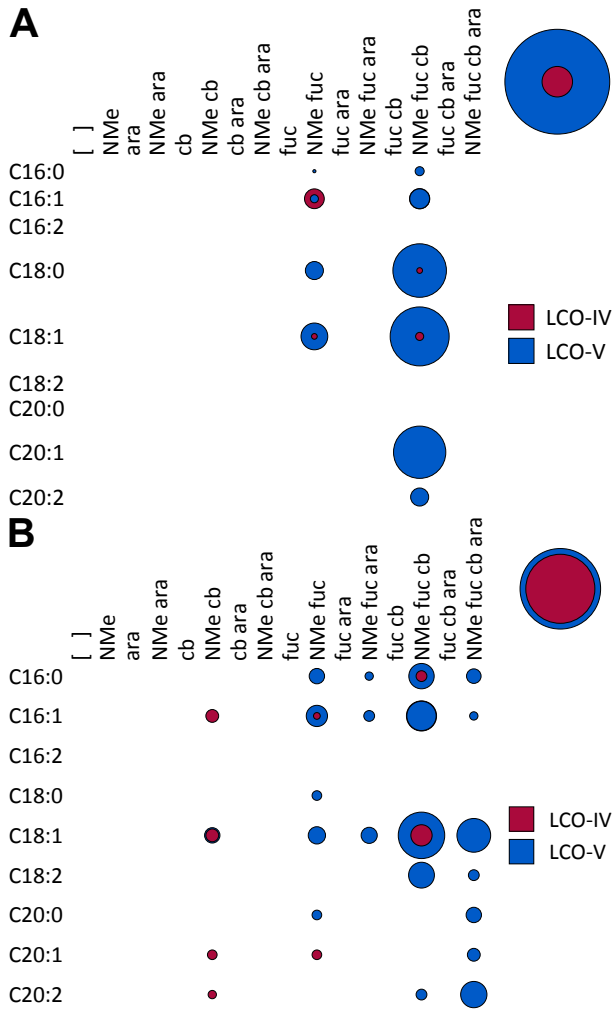
Fig. 10. A. A bubble matrix of LCOs produced by a *noeP* mutant of *Rhizobium* sp. IRBG74. Abbreviations as in Fig. 1. **B**. A bubble matrix of LCOs produced by a *noeN* mutant of *Rhizobium* sp. IRBG74. Abbreviations as in Fig. 1.

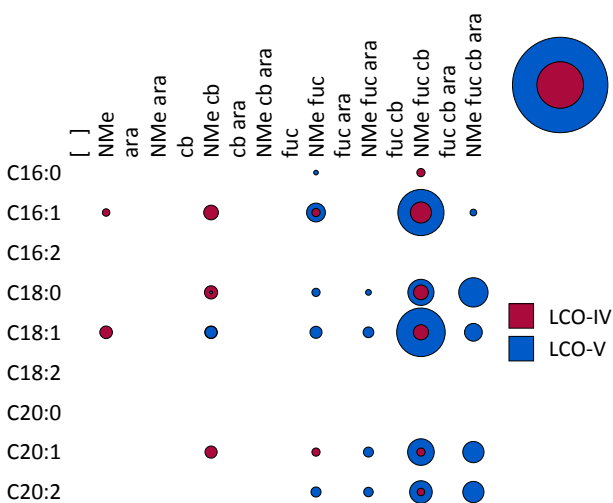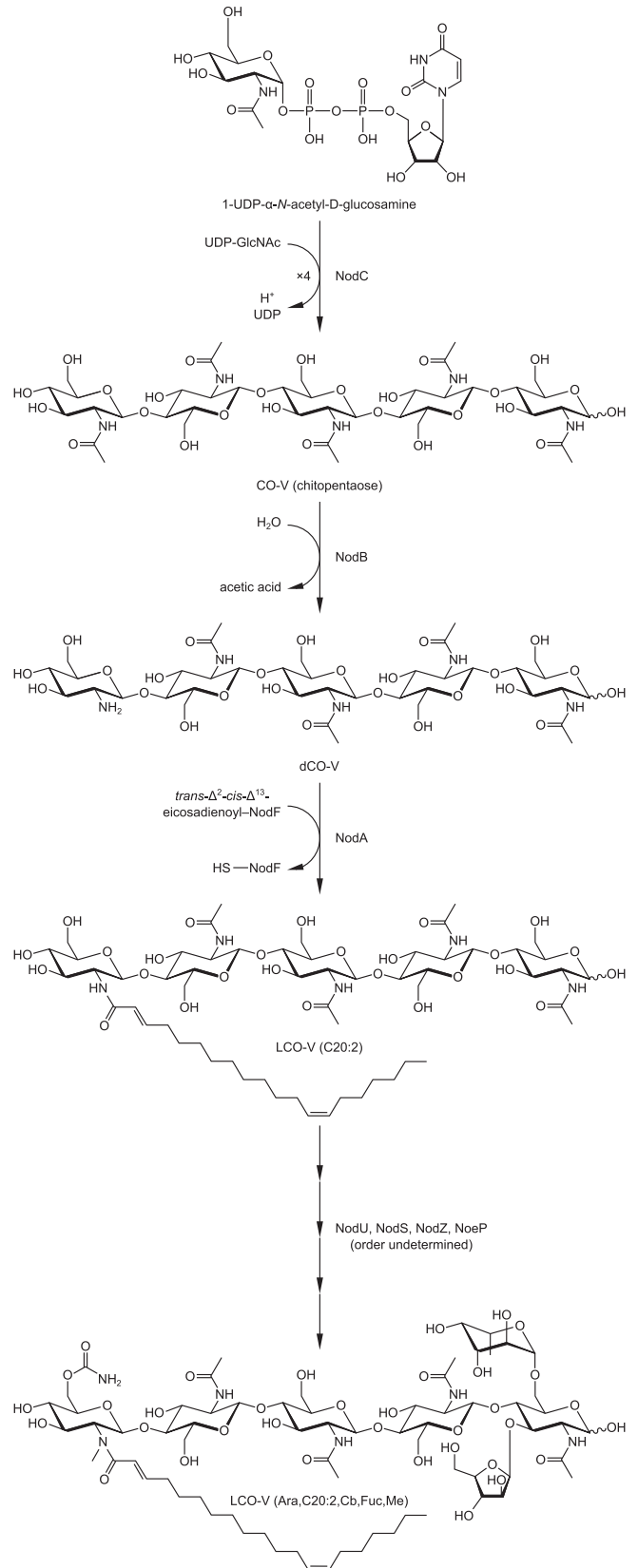**Fig. 11.** A bubble matrix of LCOs produced by an *IRBLv2_p0345* mutant of *Rhizobium* sp. IRBG74. Abbreviations as in Fig. 1.

Scheme 1. The traditionally assumed model of LCO biosynthesis, with the major LCO of *Rhizobium* sp. IRBG74 as the example. NodC, NodB, and NodA act in sequence to create the LCO backbone, followed by the addition of various decorations in no particular order.

(Scheme 2). The next step appears to be fucosylation, as evidenced by the fact that roughly half of the COs (principally as CO-IV) produced by wild-type IRB74 (Fig. 3) and the *nodB* mutant (Fig. 5) are fucosylated. This is consistent with the observations of Quinto et al. that NodZ prefers COs over LCOs as substrates [42]. CO-III (Fuc) and LCO-III (Fuc) were rarely observed, suggesting that NodZ prefers CO-IV as its substrate. Furthermore, since the percentage of LCO-V molecules declines in relation to the percentage of LCO-IV molecules in the *nodZ* mutant (Fig. 7), NodC may normally terminate the chitin chain at four residues, but fucosylation by NodZ could promote one more round of polymerization. However, this is a complicated trait since the degree of polymerization is also diminished in the *nodA*, *nodB*, and *nodS* mutants (Table 1), all of which are impaired in enzymes that act at the non-reducing end of the chitin chain, where polymerization occurs. All other additions to the chitin backbone are strictly dependent on mono-deacetylation (Fig. 5).

N-methylation occurs after mono-deacetylation (Fig. 5), as has been suggested previously [18,33]. No un-*N*-methylated compounds were detected from the wild-type strain (Fig. 2; Fig. 3). Furthermore, COs were never *N*-methylated, either in the wild-type strain (Fig. 3) or the *nodA* (Fig. 6) or *nodB* (Fig. 5) mutants. Carbamoylation by NodU is generally subsequent to N-methylation by NodS as carbamoylation is reduced (but not abolished) in the *nodS* mutant (Fig. 8) and rarely occurs independently of N-methylation in the *nodA* mutant (Fig. 6), whereas N-methylation is unaffected in the *nodU* mutant (Fig. 9). This result would suggest that NodU is adapted to use *N*-methylated dCOs as substrates but is non-specific enough that it can utilize non-*N*-methylated dCOs, albeit at lower efficiency. It should be noted that the reduced activity of NodU in the *nodS* mutant could, instead, be due to a reduction in *nodU* expression (despite our use of an in-frame deletion, which would be expected to circumvent such polar effects). However, this interpretation does not account for the phenotype in the *nodA* mutant.

Arabinosylation and acylation most likely occur at the very end of LCO biosynthesis in IRBG74. Arabinosylation clearly takes place after fucosylation since arabinosylated, non-fucosylated LCOs and COs are never detected in wild-type IRBG74 (Figs. 2 and 3) or any of its mutants (except for a small fraction seen in the *nodS* mutant). This result is in contrast to *Azorhizobium caulinodans* ORS571, where arabinosylation can readily occur independently of fucosylation [43]. Given the data, it is possible that in IRBG74 arabinosylation occurs on the fucose rather than on C-3 of the reducing terminal GlcNAc. To determine this conclusively will require enzyme treatment or linkage analysis in comparison with standard compounds. Arabinosylated COs are never detected in wild-type IRBG74 (Fig. 3) or the *nodB* mutant (Fig. 5), which would suggest that arabinosylation cannot occur unless the CO has been deacetylated. Arabinosylation is independent of carbamoylation (Fig. 9) and N-methylation (Fig. 8) and does not appear to affect those processes in turn (Fig. 10). The fact that carbamoylation, fucosylation, and methylation of dCOs increases in the *nodA* mutant suggests that NodA preferentially acylates decorated COs, positioning it later in the process of biosynthesis. The fatty acids employed by NodA for acylation of LCOs were fairly consistent across the mutant collection (with the exception of the *nodE* mutant), suggesting that the fatty acid species transferred by NodA are independent of all other LCO chemical substituents.

We have reached these conclusions by observing concomitant perturbations in the products of intact proteins when one such protein is removed from the biosynthetic pathway. This approach has been employed to interrogate the biosynthetic sequence of such diverse compounds as glycopeptide antibiotics [44,45], carbapenem antibiotics [46], carbazole alkaloids [47], indole alkaloids

[48], pipolythiodioxopiperazines [49], meroterpenoids [50], xanthones [51], prodiginines [52], fumonisins [53], and rhizobitoxine [54]. In many of these cases, the mutational analyses and spectral analyses were validated using complementary techniques. Similar inferences have been made by other authors about the order of LCO biosynthesis. Based on relative abundance of LCO species, Olsthoorn et al. opined that N-methylation of the non-reducing terminal GlcNAc is inhibited by fucosylation of the GlcNAc immediately adjacent to the non-reducing terminal GlcNAc and that the unidentified α-(1 → 3) fucosyltransferase responsible for this reaction preferentially acts on pentamers [55]. A similar preference for pentamers was also proposed for the *O*-acetyltransferase, NodX [56].

The order of LCO biosynthesis presented in Scheme 2 is the primary pathway suggested by our data. However, there is some evidence of flexibility in particular instances (such as the possibility of NodU carbamoylating dCOs prior to N-methylation, mentioned above). For example, there may be some variability in the working order of NodZ and NodB since the LCOs of wild type (~91%) and the dCOs of the *nodA* mutant (~84%) show a greater extent of fucosylation than the *nodB* mutant (~46%; see Table 1), suggesting that NodZ may have higher specificity for dCOs or that NodB may preferentially use fucosylated COs. Another point of flexibility is in the order of arabinosylation and acylation. Arabinosylation is still detected (though slightly reduced: ~17% in the *nodA* mutant vs. ~52% in WT; see Table 1) on the COs of the *nodA* mutant, suggesting that arabinosylation may not be entirely dependent on acylation. Alternatively, arabinosylated COs could be a favored substrate for NodA, resulting in efficient conversion to of these species to LCOs when *nodA* is intact. Ultimately it will be necessary to test the purified enzymes *in vitro* for their substrate specificity to resolve this question.

Two more caveats must be mentioned. First, when enzymes are removed, intermediates can build up and push a reaction in an unexpected direction due to the imbalance of reactants and products. Such a scenario would lead to erroneous conclusions about wild-type LCO biosynthesis. Second, our interpretation of the data rests on the assumption that the activity of the LCO biosynthetic enzymes is, for the most part, irreversible. However, our model (Scheme 2) must necessarily be adjusted if any of these enzymes (or some other unidentified enzyme) efficiently catalyzes the reverse reaction. For example, if NodU catalyzes both the forward and reverse reactions (carbamoylation and de-carbamoylation), then it may be that N-methylation sterically inhibits the reverse reaction, resulting in the accumulation of carbamoylated, *N*-methylated species (Figs. 2 and 3). In this scenario the *nodS* mutant would exhibit lower levels of carbamoylation because NodU is no longer inhibited from catalyzing the reverse reaction (Fig. 8). Furthermore, our model (Scheme 2) is based on the apparently preferred substrates for the LCO biosynthetic enzymes, and so it can only, at best, infer the favored biosynthetic sequence. The fact that most of the enzymes examined were still active to some degree in the various mutant backgrounds tested (the only clear exceptions being the total absence of most substitutions in the *nodB* strain and the abolishment of arabinosylation in the *nodZ* strain) suggests some degree of substrate flexibility (*i.e.* LCO biosynthesis is not a strictly stepwise, linear biosynthetic pathway). It must also be acknowledged that the diversity of CO and LCO species produced belies a certain laxity in the biosynthetic sequence. Whether our proposed model (Scheme 2) is truly the primary pathway and/or whether there is extensive cross-feeding between many possible pathways in wild-type IRBG74 (or other rhizobia) remains to be demonstrated.

Ultimately, additional analyses, such as using purified enzymes and purified substrates, heterologous expression, rescue experiments, and/or specific labeling (*e.g.* using stable isotopes), will be
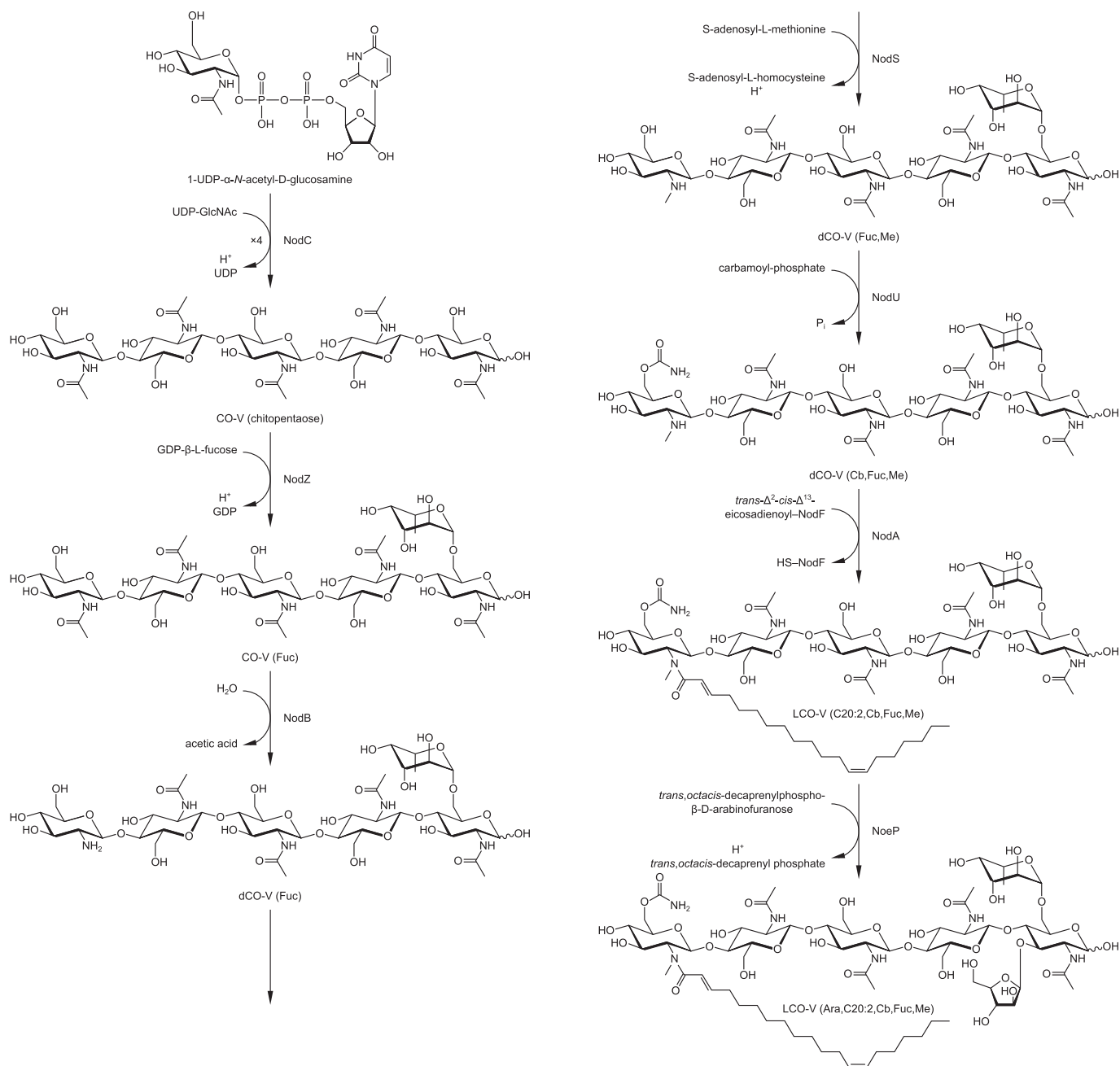
Scheme 2. An alternative model of LCO biosynthesis, as suggested by the MS data, with the major LCO of *Rhizobium* sp. IRBG74 as the example.

required to settle these questions and validate the results from the mutational analyses. For example, Quinto et al. [42] and Quesada-Vincens et al. [57] independently used a radiolabeled substrate to demonstrate the preference of NodZ for COs over LCOs. It has also previously been shown, by quantitative substrate analysis, that the preferred substrate of NodL, an *O*-acetyltransferase that acetylates C-6 of the non-reducing terminal GlcNAc, is dCOs [58–60]. On the other hand, NoeE, a sulfotransferase that sulfurylates C3 of the 2-*O*-methylfucose of fucosylated Nod factors, and NolL, an acetyl-transferase that acetylates C4 of the 2-*O*-methylfucose of fucosy-lated Nod factors, have been shown, using purified substrates, to act after fucosylation [61] and after acylation [62,63]. It has also been shown, using purified enzymes, that NodS is inactive on NodL products, but not *vice versa* [64]. Heterologous expression has been used to show that NodA can exhibit some degree of specificity in

the fatty acids it will transfer to LCOs [25,65] and that NodH and NodL are capable of, respectively, sulfating and acetylating dCOs [66]. Based on our data, it is not unreasonable to anticipate that a preference for dCOs might be shown for the majority of Nod-factor-modifying enzymes of IRBG74.

The most remarkable observation from this study is that short dCOs can be substituted with many of the same decorations as LCOs in rhizobia. This may be a result of acylation being one of the final steps (rather than one of the early steps) in the biosynthesis of LCOs. We show here that short, decorated dCOs are produced by two different species of rhizobia, *Rhizobium* sp. IRBG74 (Fig. 2; Fig. 6) and *Sinorhizobium meliloti* 1021 (Fig. S3), which suggests that it may be the case for most, if not all, rhizobia. In wild-type rhizobia the majority of LCOs assimilate into the bacterial membrane [12], but hydrophilic COs and dCOs would be capable of diffusing across

the membrane and into the extracellular milieu, where they can be readily detected by other organisms. For example, short COs have been shown to induce calcium spiking in soybean [67], *Medicago truncatula* [68], and rice [69]; to induce flavonoid production in *Medicago truncatula* [70]; to induce cortical cell divisions in vetch [71] and transgenic clover [72]; to perturb auxin flow in clover [73]; and to stimulate early nodulin expression in soybean [74]—all requisite steps in the formation of a root nodule. These observations, coupled with the data presented here, suggest that decorated COs may be hitherto unrecognized signaling molecules between plants and microbes, including rhizobia.

4. Experimental section

4.1. Microbiological techniques

4.1.1. Culture techniques

For CO and LCO production, cells were grown overnight in 60 mL of minimal medium (10 g sodium glutamate, 0.22 g K_2HPO_4, 0.1 g $MgSO_4$, 0.1 g NaCl, 20 mg $FeCl_3$, and 0.5 mg biotin per L of water) at 30 °C at 200 rpm and added to 6 L of fresh medium. After induction with 500 nM apigenin, cells were grown for 16 h at 30 °C at 200 rpm.

For strain and plasmid construction, *E. coli* and *Rhizobium* sp. IRBG74 were grown on LB agar at 37 °C and 30 °C, respectively. When necessary, the LB agar was supplemented with one or more of the following: gentamicin (Gm, 50 ng/μL for *E. coli* and 150 ng/μL for *Rhizobium*), spectinomycin (Sp, 50 ng/μL), sucrose (10%), tetracycline (Tc, 10 ng/μL).

4.1.2. Strain and plasmid construction

Bacterial strains and plasmids are described in Supplementary data, Table S1 (see also Fig. 1A). All custom primers were ordered from Integrated DNA Technologies and are described in Supplementary data, Table S2. To ensure that downstream genes were unaffected in their expression, in-frame deletions of *nod* genes were created using the method of Quandt and Hynes [75]. Briefly, approximately 500 bp of upstream and downstream sequence were amplified and spliced together by overlap-extension polymerase chain reaction using *Pfu* polymerase and then cloned into the *sacB* suicide vector, pJQ200SK [75], using Gateway® cloning (Invitrogen). The resulting deletion constructs were introduced into IRBG74 by triparental mating, followed by selection on LB-Sp-Gm. Integration of the deletion construct into the correct site was confirmed by PCR, and then de-integration was selected for by plating on LB-sucrose. Colonies that were resistant to sucrose and sensitive to gentamicin were checked by PCR for deletion of the target gene. To construct the plasmids overexpressing the *nodD* genes of IRBG74, the coding region along with the predicted ribosome binding site was amplified using *Pfu* polymerase and then cloned into the expression vector, pRF771 [76], using Gateway® cloning (Invitrogen). The resulting expression constructs were introduced into IRBG74 by triparental mating followed by selection on LB-Sp-Tc.

4.2. Purification

4.2.1. Purification of LCOs

LCOs were extracted from filtered culture supernatants by butanol extraction, as described previously [77]. Purification was first performed by HPLC on a semi-preparative C18 reverse phase column (7.5 × 250 mm, Spherisorb ODS2, 5 μm) with 10 min in isocratic solvent A (water/acetonitrile 80:20 v/v), followed by a linear gradient from solvent A to solvent B (100% acetonitrile) for 40 min at a flow rate of 1 mL min^{-1}, monitoring the UV absorption at 206 nm and 220 nm.

4.2.2. Purification of COs

For purification of COs, the butanol-extracted water phase was first concentrated on a rotative evaporator at 40 °C under 40 mBar. The concentrated phase was then freeze-dried. The powder was then resuspended at 10^{-4} M in water, sonicated, centrifuged, and placed in an O dialyzing tube for an overnight desalting with a cutoff of 1000 Da. (COs are not easily soluble in water at the concentrations used here, so they form aggregates and are not lost.) Finally, the desalted solution was set to 20% acetonitrile (ACN) and transferred to a glass vial for LC/MS analyses.

4.3. Mass spectrometry

4.3.1. MS analysis on Q-Tof

HPLC-produced fractions were analyzed using an ESI-QqToF Ultima apparatus (Waters, Milford, Massachusetts, USA) using direct infusion. Spectra were recorded in both the positive and the negative mode. Peaks detected in the expected range (m/z 1000–1500 for the single-charged species or 600–700 for the double-charged species) were submitted to MS/MS analysis, and exact masses were recorded to confirm that they were LCOs. Energies were as follows: probe, 3 kV; cone, 100 V; Rf, 70 V; collision cell, 15 V for MS and 30 V for MS/MS. Collision gas: argon. Direct inlet: solvent $AcCN/H_2O$ 1:1, 1% acetic acid, rate: 10 μL min^{-1}. Concentrations were about 10^{-4} mol L^{-1}.

4.3.2. LC/MS analyses on Q-trap

4.3.2.1. LCO analyses. For analyses of LCOs, butanol extracts or prepurified HPLC fractions were submitted to UPLC-(MS/)MS analyses. The column was an Acquity C18 BEH 150 × 2.1 mm, 1.7 μm. The gradient was 20% ACN in water with 0.1% formic acid (2 min), up to 100% ACN in 9 min, back to 20% in 1 min. The UPLC was a Dionex Ultimate 3000 system. Injected volume was 10 μl, concentrations were around 10^{-5} M. The mass spectrometer was an AB Sciex Qtrap 4500 device. Analyses were performed in the positive electrospray mode with curtain gas. The compounds were well separated and presented no matrix effect as determined by standard spiking assays. The area of the protonated monoisotopic molecular ion of each species was measured and reported. Quantification was approximate due to the limited availability of standard compounds necessary for a more precise analysis. We assume that the responses of structurally related compounds are similar.

4.3.2.1.1. EMS experiments. For enhanced mass spectrometry (EMS) experiments, the linear trap quadrupole (LTQ) was used to accumulate signal over 300 ms in order to enhance the intensity of the molecular ions detection. Then the scan occurred from 700 to 1900 m/z at a scan rate of 10,000 scans s^{-1}. Data were registered in continuum mode. Protonated molecular ions were analyzed for retentions times ranging from 5 to 8 min.

4.3.2.1.2. EPI experiments. For enhanced product ions (EPI) experiments, which provide a better detection of daughter ions, analyses were performed in enhanced product ion mode. Q1 was set to the parent ion mass and Q3 was in LTQ mode, as in the EMS experiment. Collision energy ramped from 10 to 35 V.

4.3.2.1.3. MRM experiments. For multiple reaction monitoring (MRM) experiments, which were used for the detection of minor compounds that cannot be seen directly on the ESI/MS spectra, up to 30 transitions were programmed on the system with two transitions for a single protonated molecular ion. The collision energy was optimized and ranged from 20 to 30 V, depending on the sugar backbone length. The compounds were assumed to be present when the two transitions were present at the same retention time.

4.3.2.2. CO analyses. For analyses of COs, the dialyzed fractions were subjected to HPLC-MS/MS analyses. The column was a

Hypercarb 150 × 2.1 mm, 3 μm. The gradient was 0% ACN in water with 0.1% acetic acid (1 min), up to 100% ACN in 40 min, back to 0% in 2 min. The UPLC was a Dionex Ultimate 3000 system. The injected volume was 10 μl, concentrations were around 10^{-4} M. The mass spectrometer was an AB Sciex Qtrap 4500 device. Analyses were performed in the positive electrospray mode with curtain gas. The compounds were well separated and presented no matrix effect as determined by standard spiking assays. The area of the protonated monoisotopic molecular ion of each species was measured and reported. Quantification was approximate due to the limited availability of standard compounds necessary for a more precise analysis. We assume that the responses of structurally related compounds are similar.

4.3.2.2.1. MRM experiments. For MRM experiments, up to 30 transitions were programmed on the system with two transitions for a single protonated molecular ion. The collision energy was optimized and ranged from 20 to 30 V, depending on the sugar backbone length. The compounds were assumed to be present when the two transitions were present at the same retention time.

Conflict of interest

None declared.

Acknowledgements

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.carres.2016.08.001.

References

[1] H.P. Spaink, Annu. Rev. Microbiol. 54 (2000) 257–288, http://dx.doi.org/10.1146/annurev.micro.54.1.257.

[2] J. Dénarié, F. Debellé, C. Rosenberg, Annu. Rev. Microbiol. 46 (1992) 497–531, http://dx.doi.org/10.1146/annurev.mi.46.100192.002433.

[3] J. Dénarié, F. Debellé, J.C. Promé, Annu. Rev. Biochem. 65 (1996) 503–535, http://dx.doi.org/10.1146/annurev.biochem.65.1.503.

[4] W. D'Haeze, M. Holsters, Glycobiology 12 (6) (2002) 79R–105R, http://dx.doi.org/10.1093/glycob/12.6.79R.

[5] H.P. Spaink, R.J.H. Okker, C.A. Wijffelman, E. Pees, B.J.J. Lugtenberg, Plant Mol. Biol. 9 (1) (1987) 27–39, http://dx.doi.org/10.1007/BF00017984.

[6] M.C. Peck, R.F. Fisher, S.R. Long, J. Bacteriol. 188 (15) (2006) 5417–5427, http://dx.doi.org/10.1128/JB.00376-06.

[7] W.J. Broughton, S. Jabbouri, X. Perret, J. Bacteriol. 182 (20) (2000) 5641–5652, http://dx.doi.org/10.1128/JB.182.20.5641-5652.2000.

[8] M.B. Crook, S. Mitra, J.-M. Ané, M.J. Sadowsky, P. Gyaneshwar, Genome Announc. 1 (6) (2013), http://dx.doi.org/10.1128/genomeA.00934-13 e00934–13.

[9] J.C. Biswas, J.K. Ladha, F.B. Dazzo, Soil Sci. Soc. Am. J. 64 (5) (2000) 1644, http://dx.doi.org/10.2136/sssaj2000.6451644x.

[10] J.C. Biswas, J.K. Ladha, F.B. Dazzo, Y.G. Yanni, B.G. Rolfe, Rhizobial inoculation influences seedling vigor and yield of rice, Agron. J. 92 (2000) 880–886.

[11] S.P. Cummings, P. Gyaneshwar, P. Vinuesa, F.T. Farruggia, M. Andrews, D. Humphry, G.N. Elliott, A. Nelson, C. Orr, D. Pettitt, G.R. Shah, S.R. Santos, H.B. Krishnan, D. Odee, F.M.S. Moreira, J.I. Sprent, J.P.W. Young, E.K. James, Environ. Microbiol. 11 (10) (2009) 2510–2525, http://dx.doi.org/10.1111/j.1462-2920.2009.01975.x.

[12] G.G. Orgambide, J. Lee, R.I. Hollingsworth, F.B. Dazzo, Biochemistry 34 (11) (1995) 3832–3840, http://dx.doi.org/10.1021/bi00011a041.

[13] N.P.J. Price, B. Relić, F. Talmont, A. Lewin, D. Promé, S.G. Pueppke, F. Maillet, J. Dénarié, J.-C. Promé, W.J. Broughton, Mol. Microbiol. 6 (23) (1992) 3575–3584, http://dx.doi.org/10.1111/j.1365-2958.1992.tb01793.x.

[14] S. Philip-Hollingsworth, G.G. Orgambide, J.J. Bradford, D.K. Smith, R.I. Hollingsworth, F.B. Dazzo, J. Biol. Chem. 270 (36) (1995) 20968–20977, http://dx.doi.org/10.1074/jbc.270.36.20968.

[15] E. Giraud, L. Moulin, D. Vallenet, V. Barbe, E. Cytryn, J.-C Avarre, M. Jaubert, D. Simon, F. Cartieaux, Y. Prin, G. Béna, L. Hannibal, J. Fardoux, M. Kojadinovic, L. Vuillet, A. Lajus, S. Cruveiller, Z. Rouy, S. Mangenot, B. Segurens, C. Dossat, W.L. Franck, W.-S. Chang, E. Saunders, D. Bruce, P. Richardson, P. Normand, B. Dreyfus, D. Pignol, G. Stacey, D. Emerich, A. Verméglio, C. Médigue, M. Sadowsky, Science 316 (5829) (2007) 1307–1312, http://dx.doi.org/10.1126/science.1139548.

[16] M.A. Barny, J.A. Downie, Mol. Plan Microbe Interact. 6 (5) (1993) 669–672, http://dx.doi.org/10.1094/MPMI-6-669.

[17] R.A. Geremia, P. Mergaert, D. Geelen, M. Van Montagu, M. Holsters, Proc. Natl. Acad. Sci. U. S. A. 91 (7) (1994) 2669–2673, http://dx.doi.org/10.1073/pnas.91.7.2669.

[18] P. Mergaert, W. D'Haeze, D. Geelen, D. Promé, M. Van Montagu, R. Geremia, J.C. Promé, M. Holsters, J. Biol. Chem. 270 (49) (1995) 29217–29223, http://dx.doi.org/10.1074/jbc.270.49.29217.

[19] E. Kamst, J. Pilling, L.M. Raamsdonk, B.J.J. Lugtenberg, H.P. Spaink, J. Bacteriol. 179 (7) (1997) 2103–2108.

[20] M. John, H. Röhrig, J. Schmidt, U. Wieneke, J. Schell, Proc. Natl. Acad. Sci. U. S. A. 90 (2) (1993) 625–629, http://dx.doi.org/10.1073/pnas.90.2.625.

[21] T. Ritsema, B.J.J. Lugtenberg, H.P. Spaink, J. Bacteriol. 179 (12) (1997) 4053–4055.

[22] O. Geiger, I.M. López-Lara, FEMS Microbiol. Lett. 208 (2) (2002) 153–162, http://dx.doi.org/10.1111/j.1574-6968.2002.tb11075.x.

[23] H. Röhrig, J. Schmidt, U. Wieneke, É. Kondorosi, I. Barlier, J. Schell, M. John, Proc. Natl. Acad. Sci. U. S. A. 91 (8) (1994) 3122–3126, http://dx.doi.org/10.1073/pnas.91.8.3122.

[24] E.M. Atkinson, M.M. Palcic, O. Hindsgaul, S.R. Long, Proc. Natl. Acad. Sci. U. S. A. 91 (18) (1994) 8418–8422, http://dx.doi.org/10.1073/pnas.91.18.8418.

[25] F. Debellé, C. Plazanet, P. Roche, C. Pujol, A. Savagnac, C. Rosenberg, J.C. Promé, J. Dénarié, Mol. Microbiol. 22 (2) (1996) 303–314, http://dx.doi.org/10.1046/j.1365-2958.1996.00069.x.

[26] J.P. Bourdineaud, J.J. Bono, R. Ranjeva, J.V. Cullimore, Biochem. J. 306 (Pt 1) (1995) 259–264.

[27] D.W. Ehrhardt, E.M. Atkinson, K.F. Faull, D.I. Freedberg, D.P. Sutherlin, R. Armstrong, S.R. Long, J. Bacteriol. 177 (21) (1995) 6237–6245.

[28] M. Schultze, C. Staehelin, H. Röhrig, M. John, J. Schmidt, É. Kondorosi, J. Schell, Á. Kondorosi, Proc. Natl. Acad. Sci. U. S. A. 92 (7) (1995) 2706–2709, http://dx.doi.org/10.1073/pnas.92.7.2706.

[29] N. Demont, F. Debellé, H. Aurelle, J. Dénarié, J.C. Promé, J. Biol. Chem. 268 (27) (1993) 20134–20142.

[30] G. Stacey, S. Luka, J. Sanjuan, Z. Banfalvi, A.J. Nieuwkoop, Yoon Chun Jong, L.S. Forsberg, R. Carlson, J. Bacteriol. 176 (3) (1994) 620–633.

[31] P. Mergaert, W. D'Haeze, M. Fernández-López, D. Geelen, K. Goethals, J.C. Promé, M. Van Montagu, M. Holsters, Mol. Microbiol. 21 (2) (1996) 409–419, http://dx.doi.org/10.1046/j.1365-2958.1996.6451366.x.

[32] D. Geelen, P. Mergaert, R.A. Geremia, S. Goormachtig, M. Van Montagu, M. Holsters, Mol. Microbiol. 9 (1) (1993) 145–154, http://dx.doi.org/10.1111/j.1365-2958.1993.tb01676.x.

[33] D. Geelen, B. Leyman, P. Mergaert, K. Klarskov, M. Van Montagu, R. Geremia, M. Holsters, Mol. Microbiol. 17 (2) (1995) 387–397, http://dx.doi.org/10.1111/j.1365-2958.1995.mmi_17020387.x.

[34] S. Jabbouri, R. Fellay, F. Talmont, P. Kamalaprija, U. Burger, B. Relić, J.C. Promé, W.J. Broughton, J. Biol. Chem. 270 (39) (1995) 22968–22973, http://dx.doi.org/10.1074/jbc.270.39.22968.

[35] S. Jabbouri, B. Relić, M. Hanin, P. Kamalaprija, U. Burger, D. Promé, J.C. Promé, W.J. Broughton, J. Biol. Chem. 273 (20) (1998) 12047–12055, http://dx.doi.org/10.1074/jbc.273.20.12047.

[36] H. Huang, M.S. Scherman, W. D'Haeze, D. Vereecke, M. Holsters, D.C. Crick, M.R. McNeil, J. Biol. Chem. 280 (26) (2005) 24539–24543, http://dx.doi.org/10.1074/jbc.M504068200.

[37] K. Mikušová, H. Huang, T. Yagi, M. Holsters, D. Vereecke, W. D'Haeze, M.S. Scherman, P.J. Brennan, M.R. McNeil, D.C. Crick, J. Bacteriol. 187 (23) (2005) 8020–8025, http://dx.doi.org/10.1128/JB.187.23.8020-8025.2005.

[38] G.-P. Yang, F. Debellé, A. Savagnac, M. Ferro, O. Schiltz, F. Maillet, D. Promé, M. Treilhou, C. Vialas, K. Lindström, J. Dénarié, J.-C Promé, Mol. Microbiol. 34 (2) (1999) 227–237, http://dx.doi.org/10.1046/j.1365-2958.1999.01582.x.

[39] J. Österman, J. Marsh, P.K. Laine, Z. Zeng, E. Alatalo, J.T. Sullivan, J.P.W. Young, J. Thomas-Oates, L. Paulin, K. Lindström, BMC Genomics 15 (2014) 500, http://dx.doi.org/10.1186/1471-2164-15-500.

[40] E. Ormeño-Orrillo, P. Menna, L.G.P. Almeida, F.J. Ollero, M.F. Nicolás, E. Pains Rodrigues, A. Shigueyoshi Nakatani, J.S. Silva Batista, L.M. Oliveira Chueire, R.C. Souza, A.T.R. Vasconcelos, M. Megías, M. Hungria, E. Martínez-Romero, BMC Genomics 13 (2012) 735, http://dx.doi.org/10.1186/1471-2164-13-735.

[41] M.J. Althabegoiti, E. Ormeño-Orrillo, L. Lozano, G. Torres Tejerizo, M.A. Rogel, J. Mora, E. Martínez-Romero, BMC Microbiol. 14 (2014) 6, http://dx.doi.org/10.1186/1471-2180-14-6.

[42] C. Quinto, A.H. Wijfjes, G.V. Bloemberg, L. Blok-Tip, I.M. López-Lara, B.J. Lugtenberg, J.E. Thomas-Oates, H.P. Spaink, Proc. Natl. Acad. Sci. U. S. A. 94 (9) (1997) 4336–4341, http://dx.doi.org/10.1073/pnas.94.9.4336.

[43] P. Mergaert, M. Ferro, W. D'Haeze, M. Van Montagu, M. Holsters, J.C. Promé, Mol. Plant. Microbe. Interact. 10 (5) (1997) 683–687, http://dx.doi.org/10.1094/MPMI.1997.10.5.683.

[44] D. Bischoff, S. Pelzer, B. Bister, G.J. Nicholson, S. Stockert, M. Schirle, W. Wohlleben, G. Jung, R.D. Süssmuth, Angew. Chem. Int. Ed. 40 (24) (2001) 4688–4691, http://dx.doi.org/10.1002/1521-3773(20011217)40:24<4688::

AID-ANIE4688>3.0.CO;2-M.

[45] B. Hadatsch, D. Butz, T. Schmiederer, J. Steudle, W. Wohlleben, R. Süssmuth, E. Stegmann, Chem. Biol. 14 (9) (2007) 1078–1089, http://dx.doi.org/10.1016/j.chembiol.2007.08.014.

[46] M. Rodríguez, L.E. Núñez, A.F. Braña, C. Méndez, J.A. Salas, G. Blanco, Antimicrob. Agents Chemother. 55 (4) (2011) 1638–1649, http://dx.doi.org/10.1128/AAC.01366-10.

[47] S. Huang, S.S. Elsayed, M. Lv, J. Tabudravu, M.E. Rateb, R. Gyampoh, K. Kyeremeh, R. Ebel, M. Jaspars, Z. Deng, Y. Yu, H. Deng, Chem. Biol. 22 (12) (2015) 1633–1642, http://dx.doi.org/10.1016/j.chembiol.2015.10.012.

[48] X. Gao, Y.H. Chooi, B.D. Ames, P. Wang, C.T. Walsh, Y. Tang, J. Am. Chem. Soc. 133 (8) (2011) 2729–2741, http://dx.doi.org/10.1021/ja1101085.

[49] C.J. Guo, H.H. Yeh, Y.M. Chiang, J.F. Sanchez, S.L. Chang, K.S. Bruno, C.C.C. Wang, J. Am. Chem. Soc. 135 (19) (2013) 7205–7213, http://dx.doi.org/10.1021/ja3123653.

[50] H.-C. Lo, R. Entwistle, C.-J. Guo, M. Ahuja, E. Szewczyk, J.-H. Hung, Y.-M. Chiang, B.R. Oakley, C.C.C. Wang, J. Am. Chem. Soc. 134 (10) (2012) 4709–4720, http://dx.doi.org/10.1021/ja209809t.

[51] J.F. Sanchez, R. Entwistle, J.-H. Hung, J. Yaegashi, S. Jain, Y.-M. Chiang, C.C.C. Wang, B.R. Oakley, J. Am. Chem. Soc. 133 (11) (2011) 4010–4017, http://dx.doi.org/10.1021/ja1096682.

[52] N.R. Williamson, H.T. Simonsen, R.A.A. Ahmed, G. Goldet, H. Slater, L. Woodley, F.J. Leeper, G.P.C. Salmond, Mol. Microbiol. 56 (4) (2005) 971–989, http://dx.doi.org/10.1111/j.1365-2958.2005.04602.x.

[53] R.S. Bojja, R.L. Cerny, R.H. Proctor, L. Du, J. Agric. Food Chem. 52 (10) (2004) 2855–2860, http://dx.doi.org/10.1021/jf035429z.

[54] T. Yasuta, S. Okazaki, H. Mitsui, K.I. Yuhashi, H. Ezura, K. Minamisawa, Appl. Environ. Microbiol. 67 (3–12) (2001) 4999–5009, http://dx.doi.org/10.1128/AEM.67.11.4999.

[55] M.M.A. Olsthoorn, I.M. López-Lara, B.O. Petersen, K. Bock, J. Haverkamp, H.P. Spaink, J.E. Thomas-Oates, Biochemistry 37 (25) (1998) 9024–9032, http://dx.doi.org/10.1021/bi972937r.

[56] J.L. Firmin, K.E. Wilson, R.W. Carlson, A.E. Davies, J.A. Downie, Mol. Microbiol. 10 (2) (1993) 351–360, http://dx.doi.org/10.1111/j.1365-2958.1993.tb01961.x.

[57] D. Quesada-Vincens, R. Fellay, T. Nasim, V. Viprey, U. Burger, J.C. Promé, W.J. Broughton, S. Jabbouri, J. Bacteriol. 179 (16) (1997) 5087–5093.

[58] G. Bloemberg, J.E. Thomas-Oates, B. Lugtenberg, H. Spaink, Mol. Microbiol. 11 (4) (1994) 793–804, http://dx.doi.org/10.1111/j.1365-2958.1994.tb00357.x.

[59] H.P. Spaink, A.H.M. Wijfjes, K.M.G.M. der van Drift, J. Haverkamp, J.E. Thomas-Oates, B.J.J. Lugtenberg, Mol. Microbiol. 13 (5) (1994) 821–831, http://dx.doi.org/10.1111/j.1365-2958.1994.tb00474.x.

[60] G.V. Bloemberg, R. Lagas, S. van Leeuwen, G. Van der Marel, J. Van Boom, B. Lugtenberg, H. Spaink, Biochemistry 34 (39) (1995) 12712–12720.

[61] C. Pacios Bras, M.A. Jordá, A.H. Wijfjes, M. Harteveld, N. Stuurman, J.E. Thomas-Oates, H.P. Spaink, Mol. Plan Microbe Interact. 13 (4) (2000) 475–479, http://dx.doi.org/10.1094/MPMI.2000.13.4.475.

[62] D. Quesada-Vincens, M. Hanin, W.J. Broughton, S. Jabbouri, Mol. Plant. Microbe. Interact. 11 (7) (1998) 592–600, http://dx.doi.org/10.1094/MPMI.1998.11.7.592.

[63] S. Berck, X. Perret, D. Quesada-Vincens, J.C. Promé, W.J. Broughton, S. Jabbouri, J. Bacteriol. 181 (3) (1999) 957–964.

[64] I.M. López-Lara, D. Kafetzopoulos, H.P. Spaink, J.E. Thomas-Oates, J. Bacteriol. 183 (11) (2001) 3408–3416, http://dx.doi.org/10.1128/JB.183.11.3408-3416.2001.

[65] T. Ritsema, A.H.M. Wijfjes, B.J.J. Lugtenberg, H.P. Spaink, Mol. Gen. Genet. 251 (1) (1996) 44–51, http://dx.doi.org/10.1007/BF02174343.

[66] E. Samain, V. Chazalet, R.A. Geremia, J. Biotechnol. 72 (1–2) (1999) 33–47, http://dx.doi.org/10.1016/S0168-1656(99)00048-6.

[67] J. Müller, C. Staehelin, Z.P. Xie, G. Neuhaus-Url, T. Boller, Plant Physiol. 124 (2) (2000) 733–740, http://dx.doi.org/10.1104/pp.124.2.733.

[68] A. Genre, M. Chabaud, C. Balzergue, V. Puech-Pagès, M. Novero, T. Rey, J. Fournier, S. Rochange, G. Bécard, P. Bonfante, D.G. Barker, New Phytol. 198 (1) (2013) 190–202, http://dx.doi.org/10.1111/nph.12146.

[69] J. Sun, J.B. Miller, E. Granqvist, A. Wiley-Kalil, E. Gobbato, F. Maillet, S. Cottaz, E. Samain, M. Venkateshwaran, S. Fort, R.J. Morris, J.-M. Ané, J. Dénarié, G.E.D. Oldroyd, Plant Cell 27 (3) (2015) 823–838, http://dx.doi.org/10.1105/tpc.114.131326.

[70] A. Savouré, C. Sallaud, J. El-Turk, J. Zuanazzi, P. Ratet, M. Schultze, Á. Kondorosi, R. Esnault, É. Kondorosi, Plant J. 11 (2) (1997) 277–287, http://dx.doi.org/10.1046/j.1365-313X.1997.11020277.x.

[71] H.R. Schlaman, A.A. Gisel, N.E. Quaedvlieg, G.V. Bloemberg, B.J. Lugtenberg, J.W. Kijne, I. Potrykus, H.P. Spaink, C. Sautter, Development 124 (23) (1997) 4887–4895, http://dx.doi.org/10.1016/S1369-5266(98)80007-1.

[72] C.L. Díaz, H.P. Spaink, J.W. Kijne, Mol. Plant Microbe Interact. 13 (3) (2000) 268–276, http://dx.doi.org/10.1094/MPMI.2000.13.3.268.

[73] U. Mathesius, H.R.M. Schlaman, H.P. Spaink, C. Sautter, B.G. Rolfe, M.A. Djordjevic, Plant J. 14 (1) (1998) 23–34, http://dx.doi.org/10.1046/j.1365-313X.1998.00090.x.

[74] E. Minami, H. Kouchi, J.R. Cohn, T. Ogawa, G. Stacey, Plant J. 10 (1) (1996) 23–32.

[75] J. Quandt, M.F. Hynes, Gene 127 (1) (1993) 15–21, http://dx.doi.org/10.1016/0378-1119(93)90611-6.

[76] D.H. Wells, S.R. Long, Mol. Microbiol. 43 (5) (2002) 1115–1127, http://dx.doi.org/10.1046/j.1365-2958.2002.02826.x.

[77] P. Roche, P. Lerouge, C. Ponthus, J.C. Promé, J. Biol. Chem. 266 (17) (1991) 10933–10940.