ARTICLE

OBC
www.rsc.org/obc

New access to lipo-chitooligosaccharide nodulation factors

Martin Ohsten Rasmussen,*[a] Bridget Hogg,[a] Jean-Jacques Bono,[a] Eric Samain[b] and Hugues Driguez[b]

[a] *CERMAV-FR-CNRS 2607, BP 53, F-38041 Grenoble cedex 9, France. E-mail: hugues.driguez@cermav.cnrs.fr; Fax: + (33) 476 547 203; Tel: (33) 476 037 664*

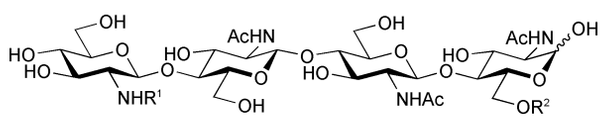
[b] *UMR 5546-CNRS-Université Paul Sabatier, 24 chemin de Borde Rouge, BP 17 Auzeville, 31326 Castanet-Tolosan, France*

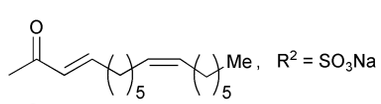
Received 9th March 2004, Accepted 29th April 2004
First published as an Advance Article on the web 9th June 2004

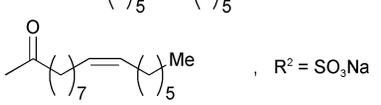
Sulfonated and non-sulfonated lipo-chitooligosaccharides involved in *Sinorhizobium meliloti*–legume symbiosis are efficiently obtained on a multi mg scale by a 2-step procedure combining biotechnological and chemical approaches.

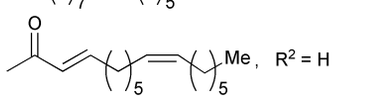
Introduction

Rhizobia are symbiotic soil bacteria that fix nitrogen in legume root nodules, through a reciprocal exchange of signal molecules. Roots of host plants secrete specific flavonoid type molecules that induce the transcription of bacterial *nod* genes that are involved in the synthesis of extracellular lipochitooligosaccharide nodulation factors (NFs). NFs signal back to the plant and set in motion the program of nodule organogenesis leading to the formation of a novel organ (the root nodule) in which the bacteria fix atmospheric nitrogen.[1] In the early nineties the first structural determination of a nodulation factor was accomplished for the major NF (**1**, Fig. 1) produced by *Sinorhizobium meliloti* (formerly *Rhizobium meliloti*) responsible for the establishment of the symbiosis with *Medicago*.[2] NFs have since been purified from various strains and species and shown to all have the same generic lipochitooligosaccharidic structure. Each bacterial strain produces NFs differing by the presence of substituents linked to the chitin backbone and/or the structure of the fatty acid moiety.



1 R^1 =  $R^2 = SO_3Na$

2 R^1 =  $R^2 = SO_3Na$

3 R^1 =  $R^2 = H$

4 $R^1 = H$, $R^2 = SO_3Na$

5 $R^1 = H$, $R^2 = H$

Fig. 1 Structures of chito- and lipo-chitooligosaccharides.

The structures of NFs appear to be at the basis of the rhizobial host specificity. Purified NFs are able to mimic many of the biological responses provoked on roots by bacteria themselves. The fact that most of these responses can be induced by a low concentration of NFs with a specific structure, strongly suggests the presence of receptors on the surface of the roots.[3–5]

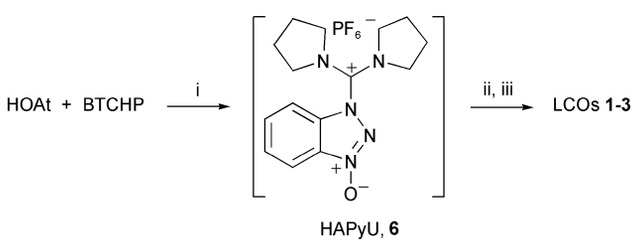
Early structure–activity relationship studies using natural NFs have been limited by the quantity produced, even by recombinant *Rhizobia* strains overproducing these compounds, and also by the difficulty in isolating pure products.[6,7] These compounds were, therefore, ideal targets for the demonstration of progress achieved in oligosaccharide synthesis by organic chemists during the last decade. The chemical synthesis of NFs of *S. meliloti* has been achieved by four different groups using four original strategies.[8–12] All these approaches led to the preparation of several lipochitotetraosyl derivatives (LCO-IV) identical to the natural NFs, and also to analogues modified on their lipidic chain (**1**–**3**, Fig. 1). Since the mid-1990s, none of these groups nor other glycochemists have wanted to repeat these time consuming syntheses and the shortage of LCOs has limited biochemical studies of *Rhizobia*–legume symbiosis. As a result of our research on developing new access to complex oligosaccharides by chemoenzymatic and biotechnological methods, we achieved the preparation of chitooligosaccharides **4** and **5** by using recombinant *Escherichia coli* strains harboring *nodC*, *nodB* and *nodH* genes coding for a chitooligosaccharide synthase, a chitooligosaccharide *N*-deacetylase and a sulfotransferase respectively.[13,14] Compounds **4** and **5** can be regarded as highly advanced intermediates for the synthesis of the complex oligosaccharide structures of nodulation factors. Indeed, the obtained chitooligosaccharides **4** and **5** have been used to prepare synthetic NFs by chemical acylation of the free amino group at the non reducing unit.[4,5,15] In this paper we describe the use of a transient activated ester for amide formation as a potent alternative to the acid chloride strategy to perform the synthesis of the main NFs produced by *S. meliloti* (compounds **1** and **2**) differing only by the structure of the fatty acid chain. The non-sulfonated homologue **3**, devoid of biological activity on *Medicago* but active on vetch, has also been synthesized using the same protocol.

Results and discussion

Approaches for oligosaccharide syntheses that use genetically engineered bacteria have emerged as alternatives to their chemical syntheses.[16] The potential of the use of living bacteria harboring heterologous genes was first demonstrated for the production of chitooligosaccharides **4** and **5**.[13,14] With the now easy availability of the advanced intermediates **4** and **5**, the selective coupling reactions between the free amine and fatty acids were taken on.

DOI: 10.1039/b403575e

Applying a classical acid chloride strategy under anhydrous conditions (DMF and DMAP) gave some amounts of *O*-acylated products (unpublished results). To avoid this undesired side reaction we then tried to carry out the reaction in aqueous medium as already described.[11] However, the acid chloride was then subjected to a considerable degree of hydrolysis, and the use of a large excess of acid chloride was necessary. With the availability of compounds **4** and **5** the access to the fatty acids can be regarded as a limiting factor – especially when multistep syntheses of these are involved, and the waste by side reactions is highly undesirable. To circumvent this obstacle we turned our attention to more modern amide coupling reactions. In recent years a considerable amount of results have been obtained in the field of liquid and solid phase synthesis of peptides, and it was tempting to explore these methods for our purpose. Close examination of the work by Albericio *et al.* on the use of different onium-based peptide coupling reactants led us to the choice of the guanidinium reagent HAPyU **6** (Scheme 1).[17] However, we have used an *in situ* formation of the activating intermediate **6**, obtained from commercially available 1-hydroxy-7-azabenzotriazole (HOAt) and *N*,*N*,*N*′,*N*′-bis(tetramethylene)-chloroformamidium hexafluorophosphate (BTCHP). Facile access to the synthetic NFs **1**–**3** could then easily be achieved after one-pot addition of the specified fatty acid followed by amine **4** or **5**. (Scheme 1) Taking into account that many higher amino sugars act as poor nucleophiles to a wide range of activated fatty acids,[18] and that the access to a family of complex oligosaccharides exemplified by **1**–**3** can now be reached in 2 steps, the yields of 32–51% seem acceptable.



Scheme 1 *Reagents and conditions*: i, DIEA, DMF, RT, 5 min; ii, $HO_2CC_{15}H_{27}$(2*E*,9*Z*) or $HO_2CC_{15}H_{29}$(9*E*), RT, 45 min; iii, **4** or **5**, RT, 4.5–6 h.

In conclusion, by combining the efficient biotechnological approach to chitooligosaccharides **4** and **5** with a one-pot chemical activation of the long chain fatty acid with **6**, rapid access to the three natural *Sinorhizobium meliloti* LCOs (**1**–**3**) has been successfully achieved. This unique strategy is now being exploited with the aim to gain access to more elaborate nodulation factors and in even larger quantities. This will facilitate the biochemical and plant physiological studies on the mechanisms involved in NF signaling during the early steps of the establishment of the *Rhizobium*–legume symbiosis and the development of potential applications in agriculture.

Experimental

NMR spectra were recorded on a Bruker AC 300 MHz spectrometer at 303 °C in either D_2O or DMSO-d_6 and the solvent residual peaks were used as internal standards, 4.71 and 2.50 ppm, respectively. *J* values are given in Hz. MALDI-TOF measurements were performed on a Bruker Daltonics Autoflex apparatus.

Silica gel column chromatography was performed on Merck Silica gel 40–63 μm with chloroform/methanol/water (70/30/5) as eluent. The sulfated LCOs were for analytical purposes further purified by HPLC on an RP18 column (Uptisphere 300 × 7.5 mm, 5 μm; Interchim, France), using a gradient of acetonitrile–aqueous K_2SO_4 (10 mM, pH = 4.6) (46/54 ⟶ 80/20).

Chitooligosaccharides **4** (CO-IV(SNa,NH_2)) and **5** (CO-IV-(NH_2)) were obtained from metabolically engineered *E. coli* as described.[13,14] Both compounds were further purified by silica gel chromatography as described earlier.[19]

LCO-IV(C16:2[2*E*,9*Z*], SNa) 1

To a solution of 1-hydroxy-7-azabenzotriazole (23 mg, 0.17 mmol) and *N*,*N*,*N*′,*N*′-bis(tetramethylene)-chloroformamidium hexafluorophosphate (37 mg, 0.11 mmol) in anhydrous DMF (0.8 mL) was added *N*,*N*-diisopropylethylamine (0.058 mL, 0.33 mmol) and the resulting solution was stirred at room temperature for 5 min, after which (2*E*,9*Z*)-hexadeca-2,9-dienoic acid[12] (28 mg, 0.11 mmol) in anhydrous DMF (0.6 mL) was added. After 45 min the above brown solution containing the activated acid was added to a solution of compound **4** (50 mg, 0.056 mmol) in anhydrous DMF (3.5 mL), and the resulting solution was stirred for 4.5 h at room temperature. The solvent was partially evaporated off, the remaining residue transferred to a large test tube, and the crude product was precipitated by addition of ethyl acetate (25 mL). Centrifugation of the mixture allowed the solid to settle and the solvents were evaporated. By a similar centrifugation procedure, the product was again washed with ethyl acetate (2 × 2 mL). Water (1 mL) was added and the remaining ethyl acetate evaporated off under vacuum. The crude product (54 mg) isolated after lyophilisation was purified by silica gel column chromatography, and following lyophilisation 33 mg (51%) of **1** was obtained as a white solid. δ_H(300 MHz, D_2O) 0.88 (3 H, m, CH_3 in chain), 1.26–1.49 (16 H, m, CH_2 in chain), 1.99–2.09 (13 H, m, 8-H, 11-H in chain and NAc), 2.17–2.24 (2 H, broad, 4-H), 3.02–3.94 (22 H, m, sugar-H), 4.07–4.26 (2 H, m, 6-H), 4.55–4.68 (3.4 H, m, β-anomeric H), 5.18 (0.6 H, d, *J* 2.1, α-anomer H), 5.39 (2 H, m, 9-H and 10-H in chain), 6.02 (1 H, d, *J* = 15.0, 2-H in chain), 6.86 (1 H, dd, *J* = 6.2 and 15.0, 3-H in chain). According to HPLC and MALDI-TOF measurements (*m*/*z* 1147.5 ($M+Na^+$)) this product contained approx. 10% of the pentameric form (+203 mass units). For further characterisation, 1 mg was submitted to semi preparative HPLC as mentioned above, and the pure LCO **1** was compared to the corresponding natural NF **1**.[6] HRMS *m*/*z* (ES^+) 1147.4443 ($M+Na^+$. $C_{46}H_{77}N_4O_{24}Na_2S$ requires 1147.4444).

LCO-IV(C16:1[9*Z*], SNa) 2

Compound **2** was prepared by using commercially available (9*E*)-hexadeca-9-enoic acid (Fluka). The protocol used was as for compound **1** (equal quantities of reagents and solvents, same order of addition and same washing and centrifugation procedures). The crude product (75 mg) isolated after lyophilisation was purified by silica gel column chromatography, and following lyophilisation 21 mg (33%) of **2** was obtained as a white solid. δ_H(300 MHz, D_2O) 0.88 (3 H, m, CH_3 in chain), 1.26–1.35 (16 H, m, CH_2 in chain), 1.61 (2 H, broad, 3-H in chain), 2.02–2.10 (13 H, m, H-8, H-11 in chain and NAc), 2.29 (2 H, broad, H-2 in chain), 3.47–3.96 (22 H, m, sugar-H), 4.05–4.22 (2 H, m, 6-H), 4.56–4.70 (3.4 H, m, β-anomeric H), 5.19 (0.6 H, m, α-anomeric H), 5.37 (2 H, m, H-9 and H10 in chain). According to HPLC and MALDI-TOF measurements (*m*/*z* 1149.5 ($M+Na^+$)), this product contained approx. 10% of the pentameric form (+203 mass units). For analytical purposes 1 mg was purified by semi preparative HPLC and pure LCO **2** was compared to the natural NF **2**. HRMS *m*/*z* (ES^+) 1149.4600 ($M+Na^+$. $C_{46}H_{79}N_4O_{24}Na_2S$ requires 1149.4600).

LCO-IV(C16:2[2*E*,9*Z*]) 3

To a solution of 1-hydroxy-7-azabenzotriazole (53 mg, 0.39 mmol) and *N*,*N*,*N*′,*N*′-bis(tetramethylene)-chloroformamidium hexafluorophosphate (86 mg, 0.26 mmol) in

anhydrous DMF (1 mL) was added *N*,*N*-diisopropylethylamine (0.136 mL, 0.78 mmol) and the resulting yellow solution was stirred at room temperature for 5 min whereafter (2*E*,9*Z*)-hexadeca-2,9-dienoic acid (66 mg, 0.26 mmol) in anhydrous DMF (0.6 mL) was added. After 45 min, the above solution containing the activated acid was added to a slightly turbid mixture of CO-IV(NH_2) **5** (85 mg, 0.11 mmol) in anhydrous DMF (5 mL) containing a spatula point of LiCl and briefly treated with ultra-sound. The resulting brown mixture turned over a few minutes into a dark brown solution. Stirring was continued at room temperature for 6 h and the solvent was partially evaporated off. The residue was transferred to a large test tube and ethyl acetate (25 mL) was added in order to precipitate the crude product. Centrifugation of the mixture allowed the solid to settle and the solvents were aspired off. By a similar procedure, the product was washed with ethyl acetate (2 × 2 mL) and water (3 × 2 mL), and the obtained white solid was lyophilised, leaving pure **3** as a white powder (35 mg, 32%). δ_H(300 MHz, DMSO-d_6) 0.86 (3 H, m, CH_3 in chain), 1.20–1.36 (14 H, m, CH_2 in chain), 1.81 (9H, m, NAc), 1.95–2.17 (6 H, m, 4-H, 8-H and 11-H in chain), 3.03–3.75 (m, sugar-H), 4.35–4.70 (3.6 H, m, β-anomeric H), 4.89 (0.6 H, broad, α-anomeric H), 5.34 (2 H, m, 9-H and 10-H in chain), 5.85 (2 H, d, J = 15.3, 2-H in chain), 6.59 (1H, dt, J = 6.6, 15.4, 3-H in chain), 7.59–7.82 (4 H, broad, NH). HRMS *m*/*z* (ES^+) 1045.5059 ($M+Na^+$. $C_{46}H_{78}N_4O_{21}Na$ requires 1045.5056).

Acknowledgements

We thank the EU for financing fellowships to M.O.R. and B.H. through the SACCSIGNET project (HPRN-CT-2002-00251). Thanks to Dr M. Rossignol (UMR UPS-CNRS 5546, Toulouse) and M. P. Jehan (CRMPO, Université de Rennes) for the MALDI-TOF and HRMS analyses respectively.

References

1 W. D'Haeze and M. Holsters, *Glycobiology*, 2002, **12**, 79R.
2 P. Lerouge, P. Roche, C. Faucher, F. Maillet, G. Truchet, J.-C. Promé and J. Dénarié, *Nature*, 1990, **344**, 781.
3 J. Dénarié, F. Debellé and J.-C. Promé, *Annu. Rev. Biochem.*, 1996, **65**, 503.
4 F. Gressent, S. Drouillard, N. Mantegazza, E. Samain, R. A. Geremia, H. Canut, A. Niebel, H. Driguez, R. Ranjeva, J. Cullimore and J. J. Bono, *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 4704.
5 F. Gressent, N. Mantegazza, J. V. Cullimore, H. Driguez, R. Ranjeva and J.-J. Bono, *Mol. Plant Microbe Interact.*, 2002, **15**, 834.
6 P. Roche, F. Debellé, F. Maillet, P. Lerouge, C. Faucher, G. Truchet, J. Dénarié and J.-C. Promé, *Cell*, 1991, **67**, 1131.
7 B. Kohring, R. Baier, K. Niehaus, A. Pühler and E. Flaschel, *Glyconjugate J.*, 1997, **14**, 963.
8 K. C. Nicolaou, N. J. Bockovich, D. R. Carcanague, C. W. Hummel and L. F. Even, *J. Am. Chem. Soc.*, 1992, **114**, 8701.
9 L.-X. Wang, C. Li, Q.-W. Wang and Y.-Z. Hui, *Tetrahedron Lett.*, 1993, **34**, 7763.
10 S. Ikeshita, A. Sakamoto, Y. Nakahara, Y. Nakahara and T. Ogawa, *Tetrahedron Lett.*, 1994, **35**, 3123.
11 D. Tailler, J.-C. Jacquinet and J. M. Beau, *J. Chem. Soc., Chem. Commun.*, 1994, 1827.
12 L.-X. Wang, C. Li, Q.-W. Wang and Y.-Z. Hui, *J. Chem. Soc., Perkin Trans. 1*, 1994, 621.
13 E. Samain, S. Drouillard, A. Heyraud, H. Driguez and R. A. Geremia, *Carbohydr. Res.*, 1997, **302**, 35.
14 E. Samain, V. Chazalet and R. Geremia, *J. Biotechnol.*, 1999, **72**, 33.
15 L. Gonzalez, M. Bernabe, J. Felix Espinosa, P. Tejero-Mateo, A. Gil-Serrano, N. Mantegazza, A. Imberty, H. Driguez and J. Jimenez-Barbero, *Carbohydr. Res.*, 1999, **318**, 10.
16 T. Endo and S. Koizumi, *Curr. Opin. Struct. Biol.*, 2000, **10**, 536.
17 F. Albericio, J. M. Bofill, A. E. Faham and S. A. Kates, *J. Org. Chem.*, 1998, **63**, 9678.
18 J. S. Debenham, R. Rodebaugh and B. Fraser-Reid, *J. Org. Chem.*, 1997, **62**, 4591.
19 P. Groves, M. O. Rasmussen, M. D. Molero, E. Samain, F. J. Caanada, H. Driguez and J. Jimenez-Barbero, *Glycobiology*, 2004, **14**, 451.