A top-down chemo-enzymatic approach towards N-acetylglucosamine-N-acetylmuramic oligosaccharides: chitosan as a reliable template

I-austo Queda‡, Gonçalo Covas‡, Tomé Silva‡, Cátia Almeida Santos‡, Maria R. Bronze§, Francisco Javier Canada†, Marta C. Corvo†, Sérgio R. Filipe*†§ and M. Manuel B. Marques*§

‡LAQV@REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Campus de Caparica, 2829-516 Caparica, Portugal

§UCIBIO@REQUIMTE, Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Campus de Caparica, 2829-516 Caparica, Portugal

†i3N|CENIMAT, Department of Materials Science, Faculty of Science and Technology, UNL, 2829-516 Caparica, Portugal

§ Laboratory of Bacterial Cell Surfaces and Pathogenesis, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, 2780-157 Oeiras, Portugal

§ iBET, Instituto de Biologia Experimental e Tecnológica, Apartado 12, 2780-901 Oeiras, Portugal

YiMED, Faculdade de Farmácia Universidade de Lisboa, Av. Prof. Gama Pinto, 1649-019 Lisboa Portugal
An unprecedented approach towards oligosaccharides containing N-acetylglucosamine-N-acetylmuramic (NAG-NAM) units was developed. These novel bacterial cell wall surrogates were obtained from chitosan via a top down approach involving both chemical and enzymatic reactions. The chemical modification of chitosan using a molecular clamp based strategy, allowed obtaining N-acetylglucosamine-N-acetylmuramic (NAG-NAM) containing oligomers. Intercalation of NAM residues was confirmed through the analysis of oligosaccharide fragments from enzymatic digestion and it was found that this route affords NAG-NAM containing oligosaccharides in 33% yield. These oligosaccharides mimic the carbohydrate basic skeleton of most bacterial cell surfaces. The oligosaccharides prepared are biologically relevant and will serve as a platform for further molecular recognition studies with different receptors and enzymes of both bacterial cell wall and innate immune system. This strategy combining both chemical modification and enzymatic digestion provides a novel and simple route for an easy access to bacterial cell wall fragments – biologically important targets.

KEYWORDS. Chitosan, chitooligosaccharides, chemo-enzymatic, regioselective modification
1. Introduction

Chitin and chitosan are biopolymers with an enormous structural potential for chemical modification (Fig.1), constituting an excellent platform for novel chemical entities with a broad scope of applications, ranging from material science to medicine.

Chitin and chitosan are polysaccharides constituted of β-(1,4) linked 2-acetamido-2-deoxy-d-glucopyranose (GlcNAc) and glucosamine (GlcN) monosaccharide repeating units, respectively (Fig. 1). The success of these biopolymers relies in their low toxicity, biocompatibility and biodegradability. (Martínez, Falomir, & Gozalbo, 2001).

[1] Due to their inherent properties, large availability and structural versatility, a great interest has been devoted to these functional biopolymers during the last years. (Miguel, Moreira, & Correia, 2019; Mincke et al., 2019; Pokhrel & Yadav, 2019; Shanmuganathan et al., 2019) Chitin and chitosan have been extensively explored, as antimicrobial agent, matrix for drug release or wound dressing material (Balan & Verestiuc, 2014; Zargar, Asghari, & Dashti, 2015), as fibers, gels, sponges, films, beads and nanoparticles, (Azuma et al., 2015) among others.

The unique structure of chitin and chitosan, when combined with a properly controlled chemical and/or enzymatic transformation, improves the design of novel materials with refined macromolecular structures and specific properties. (L. C. R. Carvalho, Queda, Santos, & Marques, 2016) The classic way to design novel materials from chitin or chitosan involves its chemical modification, and the preparation of chitin or chitosan derivatives with well-defined structures is crucial to manipulate chitin and chitosan in well-controlled manners. Undoubtedly, their chemo- and/or regioselective modification allied to a controlled molecular structure is the key for the progress on novel chitin/chitosan derivatives possessing valuable properties. (Chen et al., 2018; Kapadnis, Dey, Dandekar, & Jain, 2019; Kurita, 2001; Mourya & Inamdar, 2008)

The use of high molecular weight biopolymers to attain smaller molecules, as an alternative to classical glycosylation methods as emerged as a new breadth to attain glycostructures via a reversed synthetic approach. Indeed, chitooligosaccharides (COSs), obtained from chitin or chitosan, are highly attractive molecules for further chemical modification. (L. C. R. Carvalho et al., 2016) However COSs have been scarcely explored as potential building blocks to attain advanced and tailored glycostructures with potential application in glycobiology. (L. C. R. Carvalho, Queda,
The structural resemblance between chitin and the carbohydrate skeleton of peptidoglycan (PGN) (Fig. 1) - the major component of the bacterial cell wall, led us to recently explore chitobiose, obtained from chitosan, as starting material for PGN synthesis. (L. C. R. Carvalho et al., 2018) PGN is made of repeating \(N\)-acetylglucosamine (NAG)-\(N\)-acetylmuramic (NAM) disaccharide units, linked via [NAG-(\(\beta\)-1,4)-NAM] linkage, with a pentapeptide attached to the \(d\)-lactyl moiety of each NAM. (Holtje, 1998) The lack of pure PGN fragments, PGN derivatives or its mimetics, has strongly hampered the investigation of the relevance of the PGN structure in the metabolism of bacteria and the determination of its role in host disease. (L. H. Wang et al., 2006) The synthesis of NAG containing oligosaccharides, such as NAG-NAM disaccharides, (Berthelot et al., 2017; Shih, Chen, Cheng, Wong, & Cheng, 2011; Yang et al., 2011; Yang & Yu, 2014) presents several challenges, such as: the enantioselective establishment of the \(\beta\)-(1,4)-glycosidic bond; the use of 2-acetamido-2-deoxyglycosyl donors, that lead to formation of 1,2-O,N-oxazoline intermediates, (Arihara, Nakamura, & Hashimoto, 2005; Bongat & Demchenko, 2007; R. Enugala, Carvalho, Pires, & Marques, 2012; Stevenin, Boyer, & Beau, 2012) and the use of corresponding acceptors that are poor nucleophiles. (Crich & Dudkin, 2001) Consequently, we and others have been developing strategies for the synthesis of monomeric and dimeric muropeptides from glucosamine residues. (L. C. R. Carvalho et al., 2018; L. R. Carvalho, Corvo, Enugala, Marques, & Cabrita, 2010; R. Enugala, Carvalho, & Marques, 2010; R. Enugala & Marques, 2012; Ramu Enugala, Pires, & Marques, 2014; Fujimoto et al., 2009; Hesek, Lee, Morio, & Mobashery, 2004; Q. Q. Wang et al., 2016; Zhang et al., 2007)
Fig. 1. Structure of chitin/chitosan and of the peptidoglycan (PGN) – murein.

Since chitosan and PGN share the same basic carbohydrate backbone, and on the follow-up of our recent work on chitobiose regioselective modification, we anticipated that chitosan of high polymerization degree would be an attractive starting material for the preparation of NAG-NAM containing oligosaccharides, that can be further manipulated to construct PGN mimetics useful as probes in biological studies.(Atilano et al., 2010; Atilano, Yates, Glittenberg, Filipe, & Ligoxygakis, 2011; Kuru, Tekkam, Hall, Brun, & Van Nieuwenhze, 2015; L. H. Wang et al., 2006)

Herein we report an unprecedented and sustainable approach for the synthesis of NAG-NAM- containing oligosaccharides using raw chitosan as starting material. We envisaged that a properly functionalized, chemically modified, chitosan would be an attractive starting material for attachment of the Lac unit in the chitosan backbone via a synthetic strategy. After sequential removal of the protecting groups, the modified polymer would be recognized by hydrolytic enzymes such as lysozyme or mutanolysin(Binette & Gagnon, 2007; Goodman, Pollock, Iacono, Wong, & Shockman, 1981; Kurita, Ikeda, Yoshida, Shimojoh, & Harata, 2002; Lee et al., 2013) providing smaller oligosaccharides containing NAG-NAM units (Fig. 2). These enzymes can hydrolyze efficiently the β-(1-4) glycosidic bond in the natural PGN substrate, between NAM and NAG residues(Binette & Gagnon, 2007; Goodman et al., 1981; Kurita et al., 2002; Lee et al., 2013; Vocadlo, Davies, Laine, & Withers, 2001) and, in certain conditions and to some extent, chitin and some derivatives.(Amano & Ito, 1978; Maeda, Matsumoto, & Kondo, 1997) Furthermore, the combination of chemical and enzymatic approaches have become an extremely attractive option, relatively to the traditional orthogonal synthesis, in the synthesis of complex and biological important molecules.(Meng et al., 2014; Sardzik et al., 2012; Q. Q. Wang et al., 2016; Z. Wang et al., 2013; Wever et al., 2015)
Fig. 2. Strategy to convert chitosan of high polymerization degree into NAG-NAM containing oligosaccharides.

2. Materials and Methods

2.1. Materials

All commercially obtained reagents were used without further purification unless specified. All the mentioned solvents used in the reactions were dried by usual methods. Molecular sieves 4Å were activated by heating at 300 ºC for 3 h. Preparative and analytical TLC was performed with silica gel 60 plates of 1 mm, 0.5 mm and 0.25 mm, respectively. Nuclear Magnetic Resonance (NMR) spectra in solution were recorded at Bruker Advance 400 MHz for 1H and at 100 MHz for 13C, in CDCl₃, DMSO-d₆, D₂O or CDCl₃ with chemical shift values (δ) in ppm downfield from TMS (0 ppm) or the solvent residual peak of D₂O (4.79 ppm), DMSO-d₆ (2.50 ppm) or CDCl₃ (7.24 ppm) as internal standard. The chemical shifts (δ) for proton spectra were expressed in parts per million (ppm) and the data obtained was presented in the following order: deuterated solvent, signal chemical shift (δ), relative intensity, spin multiplicity (s – singlet, d – doublet, t – triplet, m – multiplet, dd – doublet of duplets), coupling constant (J, in Hz) and molecule peak attribution if possible. The data for carbon spectra was presented in the following order: solvent, chemical shift (δ), molecule attribution if possible. 13C CP/MAS NMR spectra were recorded both at 100.62 MHz on a Bruker Avance III 400MHz (9.4 T) spectrometer using the following parameters: time between scans(D1): 3 seconds, Channel 1H (90 MHz).
deg. pulse), (P3) pulse length: 3 microseconds (PL12), power level: 6.94 dB, Channel

$^{13}$C (P15) CP contact time: 2000 microseconds, (PL1) CP power level: 9.6 dB, 15 kHz

spinning rate; and at 75.47 MHz on a Bruker Avance III 300MHz (7.2 T) spectrometer

using the following parameters: time between scans (D1): 10 seconds, Channel 1H (90
deg. pulse) (P3) pulse length: 4 microseconds, (PL12) power level: -17.68 dB Channel

$^{13}$C (P15) CP contact time: 1200 microseconds, (PL1) CP power level: -20.79 dB,

TOSS (total suppression of spinning side bands) pulse sequence and 5 kHz spinning

rate. Infrared (IR) spectra were recorded on a Bruker Tensor 27 spectrophotometer

FTIR spectra were recorded on Perkin-Elmer Spectrum 1000 model apparatus in KBr

dispersions for solid samples or NaCl dispersions for oil samples. In each spectra

description only the more intense and characteristic bands were identified. The data

obtained is presented in the following order: sample support (NaCl or KBr); frequency

of the maximum absorption band ($v_{max}$ in cm$^{-1}$) attribution to a functional group in a

molecule if possible.

The reactions were followed by thin layer chromatography (TLC) silica gel 60

G/UV254 Macherey-Nagel with 0.20 mm. Spots detection on TLC was carried with UV

light using a 254 nm lamp (Vilber-Lourmat). Additionally, TLC plates visualization was

carried with a TLC spray solution of ethanol-sulfuric acid 9:1.

Chitosan (1) 80+ high molecular weight (MW = 393 kDa, DP = 1.800, DA = 10) was

purchased from AltaKitin. Chitosan (1') 80+ medium molecular weight (MW = 68 kDa, 

DP = 400, DA = 20) was purchased from AltaKitin. Mutanolysin from Streptomyces

globisporus was purchased from Sigma Aldrich (≥4,000 units/mg protein). Lysozyme

from egg white was purchased from Sigma Aldrich (≥40,000 units/mg protein).

2.2. Synthesis

2.2.1. Preparation of N-phthaloyl chitosan (2)

To a solution of 5.6 mmol of the anhydride – phthalic or di-phenyl maleic – in DMF (6

mL) containing 5% of water was added chitosan (1) (300 mg, 1.89 mmol), and the

mixture was heated overnight under a nitrogen atmosphere at 120 °C with stirring. The

reaction mixture was cooled and then poured into ice/water (100 mL). The precipitate

was collected on a filter, washed with 150 mL of methanol for 1 h and dried to give the
product 2 as a brown solid (400 mg) in 72% yield. $^{13}$C CP/MAS NMR (101 MHz): $\delta$

168.8 (C=O), 134.5, 130.9, 123.1 (Ar), 100.6 (C1), 82.3 (C4), 75.0 (C5), 71.0 (C3), 60.9 
(C6), 56.3 (C2). FT-IR $\nu_{max}$ (KBr): 3446, 2928, 1773, 1717, 1641, 1196-990.(Kurita et 
al., 2002)

2.2.2. Preparation of dicarboxylic acids C1 and C2

Compound C1: to a solution of phthalic anhydride (15.0 mmol) in 100 mL of anhydrous 
DCM, was added the diethylene glycol (5.0 mmol) and stirred for 10 minutes, under a 
nitrogen atmosphere. The reaction mixture was cooled to 0 °C and 15 mmol of 
anhydrous TEA was added drop-wise. The reaction was stirred overnight at room 
temperature. The solvent was evaporated, and the residue obtained cooled to 0 °C and 
200 mL of saturated sodium bicarbonate solution was added. The aqueous layer was 
was washed with ether, 3 x 100 mL. The aqueous layer was collected and cooled to 0 °C, 
acidified with diluted HCl 1 M and extracted with DCM (4 x 100 mL). (Muthusamy, 
Gnanaprakasam, & Suresh, 2006) The combined organic layers were evaporated and 
dried under reduced pressure to provide a colourless oil in 90% yield. $^1$H-NMR (400 
MHz; CDCl$_3$): $\delta$ 7.80-7.57 (m, 8H), 4.43 (t, $J = 3.4$ Hz, 4H), 3.77 (t, $J = 3.3$ Hz, 4H). 
$^{13}$C NMR (100 MHz; CDCl$_3$): $\delta$ 174.6, 167.7, 132.2, 131.8, 131.7, 131.5, 131.3, 130.8, 
129.3, 128.8, 69.1, 65.9, 65.7. FT-IR $\nu_{max}$ (NaCl): 3054, 1726, 1702, 1172-1077. HR- 
ESI: 425.0895 [M+Na$^+$] (calculated), 425.0845 (found).

Compound C2: same procedure as C1, using triethylene glycol instead of diethylene 
glycol, affording a colourless oil in 63% yield. $^1$H-NMR (400 MHz; CDCl$_3$): $\delta$ 7.81- 
7.78 (m, 4H), 7.60-7.58 (m, 4H), 4.51 (dt, $J = 4.0$, 2.2 Hz, 4H), 3.88 (dt, $J = 4.0$, 2.2 Hz, 
4H), 3.79 (s, 4H).$^{13}$C-NMR (101 MHz, CDCl$_3$): $\delta$ 171.6, 170.5, 167.7, 131.8, 131.8, 
131.5, 131.3, 131.1, 129.4, 129.2, 129.2, 129.1, 73.8, 70.8, 70.3, 69.7, 68.9, 68.9, 64.6, 
64.5. HR-ESI: 469.1111 [M+Na$^+$] (calculated), 469.1108 (found).

2.2.3. Preparation of ester compounds 3a-d from N-phthaloyl chitosan (2)

General procedure: for compound 3b: the dicarboxylic acid C1 (0.792 mmol, 0.66 
equiv.), was dissolved in 3 mL of anhydrous DMF and with CDI was added (0.792 
mmol) in, under a nitrogen atmosphere and the mixture was stirred at room temperature 
for 3 hours. The mixture was added drop-wise to a solution of 2 (300 mg, 1.2 mmol) 
and TEA (8.16 mmol) in 4 mL of anhydrous DMF, at room temperature under nitrogen
atmosphere and the resulting mixture was stirred for 24 hours. The solvent was removed
to dryness and the resulting crude washed with 5 mL of methanol in an ice bath, to
obtain 320 mg of a yellow solid. $^{13}$C CP/MAS NMR (101 MHz): $\delta$ 172.7, 168.7, 134.6,
130.4, 123.3, 100.3, 83.1, 74.7, 70.5, 65.4, 60.8, 55.9. FT-IR $\nu_{\text{max}}$ (KBr): 3428, 2924,
2860, 1771, 1387, 1289, 1196-974.

2.2.4. Preparation of silylated esters 4a-d

General procedure: for compound 4b: to a suspension of 3b (290 mg, 1.05 mmol) and
an excess of imidazole (8.53 mmol) in anhydrous DMF (40 mL) was treated with the
TBDMSCl (4.97 mmol). The reaction was stirred at room temperature for 72 hours. The
product was precipitated with 50 mL of water/ethanol (1:1 v/v) solution. Then the solid
was filtered and washed with 20 mL of water/ethanol and ethyl ether 3 x 5 mL.
Obtaining 396 mg of a yellow solid. (Kurita et al., 2002) $^{13}$C CP/MAS NMR (101
MHz): $\delta$ 172.3 (C=O amide), 168.2 (C=O ester), 148.0, 134.8, 131.8, 123.6, 116.8 (Ar),
99.2 (C1), 83.2 (C4), 82.4 (-O- CH$_2$), 75.2 (C5), 71.0 (C3), 63.4 (C6), 55.9 (C2), 25.5
(C(CH$_3$)$_3$), -6.0 (CH$_3$). FT-IR $\nu_{\text{max}}$ (KBr): 3466, 1778, 1729, 1646, 1471, 1390, 1326,
1196-970, 1119, 1071.

2.2.5. Insertion of lactate moiety - Preparation of compounds 5a-d

General Procedure: for compound 5b: To a (200 mg, 0.278 mmol) of 4b in 5 mL of
anhydrous DMF was added NaH (3.8 mmol) at 0 °C and was stirred for 1 hour. Then
the (S)-(−)-2-Chloropropionic acid was added (2 equiv.). The reaction was stirred from
0 °C to room temperature for 48 hours. Then the solvent was evaporated to dryness, the
product was diluted with 3 mL of distilled water and acidified with 1 M HCl to pH 3.
The residues were washed with ethyl ether, 3 x 5 mL, and the aqueous layers were
collected, and the solvent was evaporated. The product obtained was dried under
vacuum, obtaining 218 mg of an orange solid. FT-IR $\nu_{\text{max}}$ (KBr): 3422, 2957, 2924,
2854, 1722-1640, 1458, 1259, 1208, 1090-990, 1070.

2.2.6. Removal of protecting groups and N-acetylation - Preparation of compounds
6a-d

General Procedure: for compound 6b: to an aqueous solution of hydrazine
monohydrate (40 equiv., 6 M) was added the corresponding substrate 5b (300 mg),
and the mixture was heated to reflux for 24 h. The solvent was removed and the residue
washed with distilled water (5 mL), ethanol (5 mL) and ethyl ether (5 mL). The reaction was monitored by infra-red spectroscopy, and the product used directly in the next step without any further purification. To the previously prepared N-free chitosan derivate, acetic anhydride (20 equiv.) in dry pyridine (2 mL) at 0 ºC was added and the mixture stirred at room temperature overnight. Then the solvent was evaporated, and the crude obtained washed with ethyl ether, and to the crude obtained in 1 mL of distilled water and 0.1 mL of NaOH 1 M were added. The mixture was taken to dryness and the resulting residue dissolved in 1 mL of distilled water. The product was isolated by precipitation by addition of a solution of HCl 1 M till pH 3. The precipitated was filtered and dried. To the residue obtained a solution of TBAF 1 M in THF (2 equiv.) was added. The mixture was stirred overnight at room temperature. The solvent was removed, and the crude obtained was washed with ethyl ether (3 mL), and dissolved in distilled water (1 mL) and the pH adjusted to 3 with a solution of HCl 1 M till the product precipitated and it was filtered to give 110 mg of a light yellow solid. 6b1: 13C CP/MAS NMR (101 MHz): δ 171.9 (carboxylic acid), 171.5 (amide), 103.2 (C1), 86.3 (C4), 73.3-72.7 (C3, C5), 63.6 (C6), 55.2 (C2), 22.8 (NCH3), 20.6 (CH), 13.7 (CH3). FT-IR νmax (KBr): 3435, 1698, 1656, 1158-1012.

2.3. Enzymatic digestion
For lysozyme digestion, the samples (10 mg/mL) in MilliQ H₂O were added to 10 mg/mL of lysozyme in ammonium acetate 40 mM, pH 5.25 at a final ratio of 1:0.225 [sample (mg): enzyme (mg)] and incubated for 92 h at 37 °C and 1200 rpm in an eppendorf dry block shaker. For mutanolysin digestion the samples (10 mg/mL) in MilliQ H₂O were added to 10 mg/mL of mutanolysin in 25 mM phosphate buffer at pH 5.5 at a ratio of 1:0.15 [sample (mg): enzyme (mg)] and incubated for 92 h at 37 °C and 1200 rpm in an eppendorf heat plaque. Enzymatic digestion was stopped by heat inactivation of the enzymes. The released material was collected as the supernatant after a 5 min centrifugation at 16000 g. Reduction of the free acetal ends of the samples were carried for 2 h at room temperature by mixing an equal volume of sample and 0.5 M Borate Buffer pH 9.0 and adding freshly prepared NaBH₄ at 50 mg/mL in MilliQ H₂O in a ratio of 8.33:1 (mixture (mL): NaBH₄ (mL)). The reduction reaction was stopped by lowering the pH to 2.0 with 85% o-phosphoric acid.

Muropeptides from E coli, obtained according to the reported procedure (Boneca, Huang, Gage, & Tomasz, 2000), were digested using the procedure described above. The profile of the resulting mixture of muropeptides was acquired by HPLC separation on a reverse phase C-18 Hypersil column as described. (Boneca et al., 2000) The digested sample was separated on wide pore Nucleosil C18 reverse-phase HPLC column using a 0–10.5% acetonitrile convex gradient in 100 mM sodium phosphate buffer, pH 2, for 90 min followed by a 30% acetonitrile step in the same buffer. Peak detection was performed using a UV detector at 205 nm. The PGN fragments were obtained according to described procedure. (F. Carvalho et al., 2015)

Monitoring of the enzymatic digestion was carried by running the sample on the HPLC. A non-digested sample subjected to the same incubation conditions (time, temperature and buffer) was used as negative control for the digested sample. HPLC analysis (see Figures S21-25) best enzymatic digestion was achieved for sample 6b1, which was further analysed by LC-MS (Figure S26). The efficiency of the enzymatic digestion was determined by careful analysis of LC-MS results (see Mass Spectrometry Section).

2.4. Monosaccharide composition analysis

The monosaccharide composition of the samples was determined by HPAEC-PAD as described. (Guan & Mariuzza, 2007) 0.2 mg of sample was hydrolysed to
monosaccharides by a 2 h incubation at 95 °C in a 3 M HCl solution. The hydrolysis reaction was stopped by lyophilizing the HCl solution until dryness. The remaining HCl, which had not been evaporated, was diluted by adding 500 µL water and the resulting sample lyophilized until dryness. The released monosaccharides were resuspended on a final volume of 150 µL of MilliQ H2O. Unless stated otherwise, 10 µL of the resulting samples were analysed on a CarboPac PA10 column following a previously described method. (Covas, Vaz, Henriques, Pinho & Filipe, 2016)

HPLC data: In order to test whether chitooligosaccharides released from unmodified chitin could be present in the digestion of synthetic sample, a control experiment was performed in which chitin was subjected to the same digestion conditions. Moreover, this control experiment allowed to investigate the selectivity of the enzymes for the modified sample.

According to the literature, lysozyme digestion of partially deacetylated chitin furnishes 33.1% of monosaccharide (N-acetylglucosamine) together with 26.3% of NAG-NAG and small amounts of higher oligosaccharides. (Amano & Ito, 1978)

Next, we investigated the digestion, under the same conditions, of PGN from *E. coli*. The PGN from *E. coli* (L. H. Wang et al., 2006) was used as a model to establish the optimum conditions for digestion as well as to test the HPLC program, column and conditions to follow the digestion of synthetic samples. These PGN samples contained the oligosaccharides and the peptide stems. The HPLC conditions were adapted from a literature procedure reported for PGN from *S. aureus*. (Boneca et al., 2000)

**2.5. Mass spectra**

The LC–MS/MS analysis were performed on a Waters Alliance HPLC system (Waters, 2695 separation module, Ireland) comprising a quaternary pump, an on-line solvent degasser, autosampler and column oven. The separation of the compounds was done on a reversed-phase column (LiChrospher® 100 RP-18, 250x4 mm; 5 µm; Merck®) at 35 °C using an injection volume of 20 µL. The mobile phase consisted of a Milli-Q water containing 0.5% formic acid (A): Acetonitrile (B). A flow rate of 0.30 mL/min was used, and the gradient conditions applied consisted of a linear increase from 100% to 0% (A) in 30 min; 20 min at 0% (A) and return to 100 % (A) in 10 min. The system was
then re-equilibrated with 100% (A) for 20 min. Tandem mass spectrometry (MS/MS) detection was performed on a Micromass® Quattro Micro triple quadrupole (Waters®, Ireland) using an electrospray ionization (ESI) source operating at 120 ºC and applying a capillary voltage of 3.0 kV and cone voltage 30.0 V. High purity nitrogen (N₂) was used both as drying gas and as a nebulizing gas. Ultra high-purity argon (Ar) was used as collision gas and collision energy was 20-30 eV. MassLynx software (version 4.1) was used to control the system, for data acquisition and processing.

2.6. Quantification studies

In order to quantify the amount of each sample digested, the peak observed at 870 m/z from the TCI of the non-digested sample was used as reference (corresponding to the non-digested polymer). This peak was also observed in the LC-MS of sample 6b1, corresponding to the non-digested polymer (45% of the initial). Using the calibration curve (see Supporting Information) was possible to calculate the amount of disaccharide present in the digested polymer. The NAG-NAM disaccharide was synthesized according to the reported procedure.(R. Enugala et al., 2010)

3. Results and Discussion

3.1. Chemical modification of chitosan

The most challenging step of the synthetic sequence consisted on the establishment of the conditions for introduction of a Lac moiety at O-3 position of chitosan, in alternate units of glucosamine (Fig. 2 and Scheme 1). Chitosan (1) was used instead of chitin due to its higher solubility in organic solvents.(Pillai, Paul, & Sharma, 2009) To ensure this arrangement, a chemical strategy was designed involving a regioselective protection, and the selective attachment of the Lac at a specific O-3 position was attained by regio and steric discrimination. We anticipated that the natural orientation of the biopolymer(Pillai et al., 2009) would contribute to the differentiation between alternate glucosamine units of a derivative properly functionalized. Due to the high molecular weight of the biopolymer, the control over the stepwise modification was performed by cross polarization magic angle spinning nuclear magnetic resonance (CP/MAS NMR), infrared (FT-IR) and by ionic chromatography Dionex using NAG and NAM as control (see SI). However, the line broadening in high molecular weight chitosan
derivatives in the NMR spectra, led us to perform the same modifications with chitosan of medium molecular weight. This allowed us to better monitor the chemical modification by CP/MAS NMR.

The first step consisted on the N-protection using phthaloyl group as the N-protecting group of the commercial chitosan (1) (DP (degree of polymerization) = 1800 and DA (degree of acetylation) = 10%, DD (deacetylation degree) = 90%) (Binette & Gagnon, 2007; Kurita et al., 2002). The reaction was performed according to a reported procedure (Kurita et al., 2002). This N-protection strategy prevents the problems associated with the poor nucleophilicity of the N-acetyl glucosamine derivatives (Crich & Dudkin, 2001; R. Enugala et al., 2012). In order to promote the attachment of the Lac unit in alternate glucosamine units we next explored the use of a molecular clamp strategy, consisting on the formation of a stable bridge between two 6-OH of alternate glucosamine units, via an ester linkage. We hypothesized that this step would be able to create different accessibilities to the 3-OH of adjacent glucosamine units (red/green circles in Scheme 1). The use of an intramolecular stable bridge or molecular clamps has already been reported by Fukase and Shibata to control the stereo and enantioselectivity of glycosylations. (Wakao, Fukase, & Kusumoto, 2002)
Scheme 1. Synthesis of NAG-NAM oligosaccharides. Red/green circles highlight the 3-OH of adjacent glucosamine units. (a) Phthalic anhydride, DMF 5% H₂O, 120 °C, overnight; (b) C1 (m = 1) or C2 (m = 2), CDI, DMF, r.t., 24 h; (c) TBDMSCl, imidazole, DMF, N₂, r.t., 72 h; (d) (S)-2-CPA, NaH, DMF, r.t., 48 h; (e) i) NH₂NH₂, H₂O, reflux, 24 h; ii) Ac₂O, pyridine, r.t., overnight; (f) TBAF, pyridine, r.t., overnight;

To install the bridges between alternate glucosamines, two clamps were synthesized possessing different lengths. Thus, the dicarboxylic acids C1 and C2 were prepared according to a reported procedure. (Muthusamy et al., 2006) Preliminary studies were undertaken to establish the optimized conditions for O-benzoylation such as solvent, reaction time and concentration, to avoid cross-linking between chains and promote the intra-molecular coupling. Additionally, different amounts of the dicarboxylic acids C1 and C2 were tested to maximize the attachment of Lac moiety into compound 3 in alternate positions. Thus, treatment of 2 with pre-activated acids C1 or C2 with CDI in DMF, led to the formation of compounds 3a-d. Compounds 3a-d were further silylated at the remaining free 6-OH by reaction with TBDMSCl in the presence of imidazole, in DMF. [15b] Reports on chitosan modification have demonstrated that protection
with TDBMScI is regioselective for O-6 (degree of substitution = 1.0). (Binette & Gagnon, 2007; Goodman et al., 1981; Kurita et al., 2002)

The next step consisted on the introduction of the Lac moiety, by treatment of compounds 4a-d with NaH in DMF, using (S)-(-)-2-chloropropionic acid ((S)-2-CPA), under different conditions and equivalents (see SI, Table 1, entries 1-5). The last steps consisted on sequential removal of the phthalimide group by treatment with hydrazine solution, with simultaneous removal of the ester clamp, followed by N-acetylation with acetic anhydride to afford compounds 5a-d, and finally silyl groups removal, by treatment with a TBAF solution, afforded compounds 6a-d. Due to the polymeric character of the substrates, the structural characterization was performed by FT-IR (Fig. 3) and $^{13}$C CP/MAS NMR (Fig. 4). At this stage the complexity of the products turned the characterization quite challenging, so a parallel synthesis using a medium molecular weight chitosan (1') (DP = 398, MW = 65 kDa), was performed applying the same reaction conditions (see SI).

![Fig. 3. FT-IR comparison between high and medium molecular weight chitosan, 1 and 1', respectively: (A) N-phthaloyl chitosan (2 and 2'); (B) N-phthaloyl O-6-]
benzoyl chitosan (with C1) - 3b and 3\''; (C) N-phthaloyl O-6-benzoyl O-6-
TBDMS chitosan - 4b and 4\''; (D) with the lactyl group at O-3 – 5b1 and 5\''; (E)
after protecting groups removal – 6b1 and 6.

Fig. 3A shows the FT-IR spectra obtained for the N-phthaloyl chitosan (2) which
is in agreement with the reported. (Kurita et al., 2002) The FT-IR analysis
confirmed the N-protection by the appearance of a double band at 1773 and 1717
cm\(^{-1}\). Additionally, in the \(^{13}\)C CP/MAS spectra (Fig. 4A), the characteristic
signals in the aromatic region at 120-140 ppm were observed in both MW
substrates 2 and 2\''.

The FT-IR spectra of the product obtained from O-6 benzoylation of 2 with the
acid C1, compound 3b, shows a broad band at 1770 cm\(^{-1}\), indicating the presence
of an ester group in both MW substrates (Fig. 3B in comparison with Fig. 3A).
The \(^{13}\)C CP/MAS spectra (Fig. 4B) shows the prevalence of N-phthaloyl group,
but the introduction of the clamp in the high MW substrate, which is observed at
55-70 ppm is not so clear, probably due to the relaxation of this macromolecule.
However, in the medium MW substrate 3\'' the extra resonance due to the -OCH\(_2\)-
groups at 55-70 ppm is clearly perceived. As mentioned above, the remaining O-
6 groups were protected with TBDMS group, and the FT-IR spectra of compound
4b (Fig. 3C), shows the appearance of the band at 840 cm\(^{-1}\) and the sharpness of
the band at 3400 cm\(^{-1}\) (O-H bond). In the \(^{13}\)C CP/MAS spectra, the presence of
the TBDMS group (Kurita et al., 2002) was confirmed by the presence of the
methyl groups at -8, 19 and 25 ppm (Fig. 4C). The transformations performed on
the medium MW chitosan 4\'' allowed a clear identification of the structural
modifications with this synthetic procedure.

The FT-IR spectra of the product 5b1, obtained after lactate insertion, shows a
modification in the carbonyl region at 1643 cm\(^{-1}\) (C=O) and a large band at 3422
cm\(^{-1}\) corresponding to the free OH (COOH) (Fig. 3D). The insertion of Lac unit
at O-3 was further evaluated by the enzymatic and biological assays. The FT-IR
spectra of product 6b1 obtained after removal of the protecting groups and N-
acetylation, showed the presence of the free O-H (3445 cm\(^{-1}\)), the amide group
(1698 cm\(^{-1}\)) and a carboxylic acid group (1656 cm\(^{-1}\)), indicating that all protecting
groups have been successfully removed (Fig. 3E).

Fig. 4. $^{13}$C CP/MAS comparison between high and medium molecular weight chitosan, 1 and 1', respectively: (A) N-phthaloyl chitosan (2 and 2'); (B) N-phthaloyl O-6-benzoyl chitosan (with C1) - 3b and 3'; (C) N-phthaloyl O-6-benzoyl O-6-TBDMS chitosan - 4b and 4'.

After, structural characterization and in order to evaluate the amount of lactate moiety successfully installed, the ratio of NAM and NAG moieties was determined. The resulting compounds 6a-d were hydrolysed with HCl 3 M for 3 hours at 95 °C, followed by ionic chromatography (Dionex system), (Kaiser & Benner, 2000; Kusumoto, Imoto, Ogiku, & Shiba, 1986; Wakao et al., 2002) in order to evaluate the ratio of NAG and NAM moieties (Table 1).

The ratio NAM/NAG of bacterial PGN from Staphylococcus aureus and Escherichia coli by chromatography is 1.2 and 1, respectively, and were used as positive controls (Table 1, entries 6 and 7). First experiments were carried with both dicarboxylic acids C1 and C2 at 33% mol ratio with monomer unit of 2 and 2 equivalents of (S)-2-CPA (Table 1, entries 1 and 4). Compound 6a, prepared with C1 had a significant but lower NAM/NAG ratio (0.95) than those obtained with bacterial samples. In order to modulate the ratio NAM/NAG, a higher amount of the dicarboxylic acid was also used (Table 1, compounds 6b1 and 6d, entries 2 and 5). This resulted in a higher formation of NAM as it can be seen in compound 6b1 (1.21), which results from the use of the shorter bridge C1 at 66 mol%. Similar result was obtained for compound 6d (1.38), which results from the use of longer bridge C2. An increased amount of (S)-2-CPA also led to a higher amount of NAM as expected (compound 6b2 (1.85)). The ratio of NAM/NAG indicated a significant incorporation of NAM residues on the
chitosan skeleton, but gave no clue about how alternate the NAM units had been
installed.

**Table 1.** Strategies of chemical modification of chitosan – NAM/NAG ratio of 6a-d.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Sample</th>
<th>Bridge</th>
<th>2-CPA</th>
<th>Ratio&lt;sup&gt;[ii]&lt;/sup&gt; NAM/NAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6a</td>
<td>C1 (33)</td>
<td>2</td>
<td>0.95</td>
</tr>
<tr>
<td>2</td>
<td>6b1</td>
<td>C1 (66)</td>
<td>2</td>
<td>1.21</td>
</tr>
<tr>
<td>3</td>
<td>6b2</td>
<td>C1 (66)</td>
<td>4</td>
<td>1.85</td>
</tr>
<tr>
<td>4</td>
<td>6c</td>
<td>C2 (33)</td>
<td>2</td>
<td>1.52</td>
</tr>
<tr>
<td>5</td>
<td>6d</td>
<td>C2 (66)</td>
<td>2</td>
<td>1.38</td>
</tr>
<tr>
<td>6</td>
<td>S. aureus</td>
<td>_</td>
<td>_</td>
<td>1.2</td>
</tr>
<tr>
<td>7</td>
<td>E. coli</td>
<td>_</td>
<td>_</td>
<td>1</td>
</tr>
</tbody>
</table>

<sup>[i] 1) TBDMSCl, imidazole, DMF, N<sub>2</sub>, r.t., 72 h; 2) (S)-2-CPA (2 equiv.), NaH, DMF, r.t., 48 h; 3) NH₂NH₂, H₂O, reflux, 24 h; 4) Ac₂O, pyridine, r.t., overnight; 4) TBAF, pyridine, r.t., overnight; [ii] Measured by HPAEC-PAD using in a Dionex ICS-5000 system with a CarboPac PA10 column. Eluted with an isocratic step at a 18 mM NaOH for the first 20 min., followed by a gradient step from 20 to 25 min. of 0 to 0.2 M of sodium acetate, followed by an isocratic step from 25 to 30 min. at 0.2 M of acetate and a final gradient step from 30 to 35 min. from 0.2 to 0.8 M of acetate to wash the column;"
3.2. Enzymatic steps

The ratio of NAM/NAG indicated a significant incorporation of NAM residues on the chitosan skeleton (Table 1), but gave no clue about how alternate the NAM units had been installed. In order to evaluate the intercalation of NAM residues, samples 6a-d were further digested with hydrolytic enzymes, such as lysozyme and mutanolysin, as lytic enzymes for the natural PGN, and run on HPLC. Therefore, and with regard to release NAG-NAM containing fragments, we also used mutanolysin, which is described to be more selective for the recognition of NAG-NAM moiety on bacteria cell wall than lysozyme. (Binette & Gagnon, 2007; Goodman et al., 1981; Kurita et al., 2002; Lee et al., 2013)

Several conditions (eluents and programs) were tested in order to optimize the HPLC profile. (Boneca et al., 2000) A comparison with the digestion of PGN of *E. coli* by mutanolysin and by lysozyme was performed (see SI). However, the HPLC profile of the PGN from *E. coli* also displays the peaks from the peptide stems, additionally to the carbohydrate moiety, which complicates the chromatogram analysis.

Compound 6b1 gave the best hydrolysis profile on HPLC when digested either by mutanolysin or by lysozyme. Additionally, the ratio NAM/NAG (1.21) obtained for 6b1 by Dionex was also the most similar to the bacterial ratio of *S. aureus*. The HPLC results from enzymatic digestion of compounds 6a, 6b2, 6c and 6d (see SI) (lysozyme and mutanolysin) demonstrate that despite of the Dionex results, the relative position of NAM is not as alternated as in 6b1. Indeed, a less efficient digestion was observed for the compounds 6a, 6b2, 6c and 6d.

Compound 6b1 gave the best hydrolysis profile on HPLC when digested either by mutanolysin or by lysozyme. As shown in Fig. 5, the digestion of 6b1 by lysozyme (Fig. 5B) as well as mutanolysin (Fig. 5D) generates several fragments (containing NAG-NAM moieties) in the region 7 to 12.5 min, characteristic of the PGN carbohydrate digestion by the same enzymes (and absent in the control samples). (21) Additionally, the ratio NAM/NAG (1.21) obtained for 6b1 by Dionex was also the most similar to the bacterial ratio of *S. aureus* (Table 1). The HPLC results from enzymatic
digestion of compounds 6a, 6b2, 6c and 6d (see SI) (lysozyme and mutanolysin) demonstrate that the relative position of NAM is not as alternated as in 6b1 (Fig. 5). The weak recognition of samples 6a, 6b2, 6c and 6d by these enzymes indicates that the oligosaccharide’s composition does not mimic the PGN carbohydrate backbone (despite of the NAG/NAM ratio determined by the dionex experiments). The enzymatic digestion revealed that alternating NAM formation on chitosan backbone was obtained when a higher loading of a shorter and less flexible bridge C1 is used. (Goodman et al., 1981; Vocadlo et al., 2001) We can not exclude the modification of reaction conditions that may improve the composition of the released NAG-NAM containing saccharides. (Boneca et al., 2000)

Fig. 5. A – Lysozyme negative control (under the same conditions: 6b1, buffer, temperature and reaction time without enzyme) – black; B – Lysozyme digestion – purple; C – Mutanolysin negative control (under the same conditions: 6b1, buffer, temperature and reaction time without enzyme) – blue; D – Mutanolysin digestion – brown.

3.3. LC-MS
We next evaluated the composition of the fragments released by digestion with mutanolysin by LC-MS/MS analysis. Mutanolysin only recognizes NAG-NAM sequences, (Lee et al., 2013) hydrolyzing the glycosidic bond between NAM and NAG (does not hydrolyse chitin, Fig. S19). The m/z 499 ([NAG-NAM+H]^+ or [NAM-NAG+H]^+) and m/z 521 ([NAG-NAM+Na]^+ or [NAM-NAG+Na]^+) were selected to have a quick elucidation for the target disaccharide (Fig. 6, panels A and B). The data obtained showed that the [NAG-NAM+Na]^+ disaccharide is present with a retention time of 9.5 min (Fig. S26 and S28). The mass spectra of LC-MS/MS of disaccharide released from *E. coli* peptidoglycan and sample 6b1 showed a similar fragmentation pattern (Fig. S29). However, the most intense fragment detected on the LC-MS was observed at 10.48 min, with m/z of 724, which corresponds to the trisaccharide ([NAG-NAG-NAM+Na]^+) (Fig. 6, panel
Fig. 6. LC-MS spectra at m/z 521: E. coli (A) and 6b1 at 9.52 min. (B); MS spectra from sample 6b1 at 10.48 min. (C).

From the TCI it was possible to determine the percentage of digested polymer, and it was found that 55% of the initial 6b1 sample was solubilized by mutanolysin (see SI). The data collected by the dionex chromatography (Table 1, entry 3) indicated that 6c is composed of 1.21 NAM/NAG. Indeed, in the TCI of LC-MS analysis a peak was found at 35 to 55 min that corresponds to NAM polymer 870 m/z (not digested). Quantification studies performed by mass spectrometry indicate that from the 55% digested sample 6b1, 33% contained NAG-NAM oligosaccharides (for a reference reaction see SI). This percentage includes both the free disaccharide peak as well as the disaccharide resultant from fragmentation of higher oligosaccharides: 7% of disaccharide NAG-NAM and 26% of NAG-NAG-NAM (in separated peaks) (see Fig. S28).

The LC-MS/MS (Fig. S29) analysis combined with the data from the ionic chromatography suggests that the strategy applied favoured the 2-CPA insertion at every third glucosamine unit, distinguishing two adjacent glucosamine units. Thus, compound 6b1 (with C1) is converted into a derivative possessing NAG-NAM-NAG oligosaccharides which is recognized by mutanolysin, originating the trisaccharide NAG-NAG-NAM as major fragment along with the NAG-NAM disaccharide. Additionally, the non-hydrolysed polymer might correspond to sequences of NAM residues that were not recognized by the enzymes. This excess of NAM was observed in the ratio NAM/NAG determined.

4. Conclusions
We have developed an innovative, fast and simple strategy for the synthesis of NAG-NAM containing oligosaccharides, based on a top-down approach, combining a synthetic and an enzymatic route. With the choice of a proper molecular bridge and reaction conditions, it is possible to convert high molecular weight chitosan into valuable compounds that have been prepared until now by long multistep procedures. The oligosaccharides produced $6b1$ consist of surrogates of the carbohydrate backbone of bacterial cell wall PGN, that once hydrolysed by lytic enzymes release shorter oligosaccharides containing NAG-NAM units. The methodology we have presented can be applied to commercial chitosan and used to formation of NAG-NAM containing oligosaccharides after enzymatic digestion and separation by column chromatography. The compounds generated by this approach can be further explored on the assembly of bacterial PGN and used to investigate the recognition of bacteria by different hosts. Similar strategies can be investigated for the synthesis of other carbohydrates.

Acknowledgments

This work was funded by Fundação para a Ciência e Tecnologia (FC&T) for: fellowships SFRH/BD/89518/2012, SFRH/BD/52207/2013 (to FQ and GC); projects (PTDC/QEQ-QOR/2132/2012 and UID/DTP/04138/2013). This work was supported by the Associate Laboratory for Green Chemistry- LAQV which is financed by national funds from FCT/MCTES (UID/QUI/50006/2019) and co-financed by the ERDF under the PT2020 Partnership Agreement (POCI-01-0145-FEDER - 007265). i3N|Cenitmat is financed by national funds from FC&T/MEC (UID/CTM/50025/2019) and co-financed by the ERDF under the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007688) The National NMR Facility is supported by FC&T (ROTEIRO/0031/2013 – PINFRA/22161/2016, co-financed by FEDER through COMPETE 2020, POCI, and PORL and FC&T through PIDDAC). iNOVA4 Health Research Unit (LISBOA-01-0145-FEDER-007344), is co-funded by FC&T/Ministério da Ciência e do Ensino Superior, and by FEDER under the PT2020 Partnership Agreement. CTQ2012- 32025 and CTQ2015-64597-C2 of Spanish Ministry of Economy and Competitiveness to FJC are acknowledge. The LC-MS/MS equipment is part of The National MS Facility, supported by FC&T(REDE/1518/REM/2005).
References


Supplementary data
Click here to download Supplementary data: CarbohydrPolym_Supporting Info_Final.docx