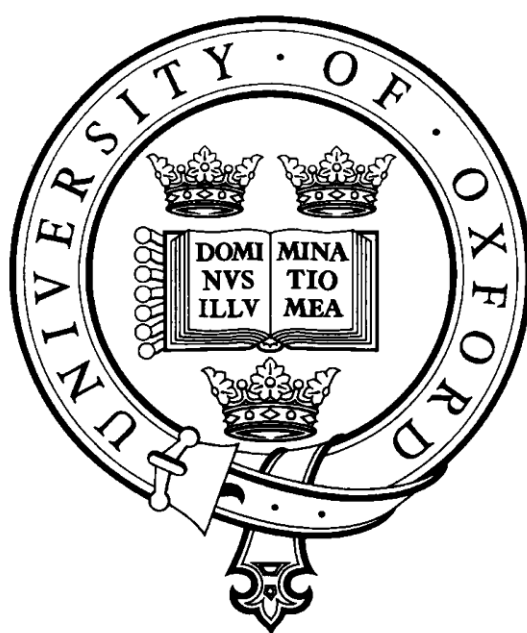


Biological investigation of glycosylated macromolecular structures



A thesis submitted to the Board of the Faculty of Physical Sciences in partial
fulfilment of the requirement for the degree of Doctor of Philosophy at the
University of Oxford

Sonia De Munari

Merton College, University of Oxford

Michaelmas 2015

Dedicated to my grandmother Dosolina

*Da soli si puo' arrivare lontano, ma per sentirsi davvero liberi
bisogna avere radici.*

Abstract

Carbohydrates play a key role in many biological processes. The complex pool of information that these structures are able to express forms the glycode. Our understanding of this language is limited, due to the complexity of the sugar structures and the diversity of their functions. The work presented in this thesis aims to contribute to our knowledge of their biological significance by building what we will call a Glycomap, a selection of biologically relevant sugars specifically picked to map the *in vivo* interactions of carbohydrates and explore some of their applications to different macromolecules: proteins, nanoparticles and carbon nanotubes.

The controlled glycosylation of proteins still represents a challenge, hence we decided to exploit the ability to diversify carbohydrates from a common precursor, and developed three glycosylating reagents which allowed the use of different glycosylation strategies on different protein structures.

In order to follow the *in vivo* distribution of the selected carbohydrates, we activated them into the 2-imino-2-methoxyethyl derivatives and used them to decorate amino-functionalised magnetic nanoparticles. These nanoparticles were isolated from different tissues after testing in animal models, and their distribution was recorded.

As sugars can also be applied to drug delivery systems as targeting agents, we explored the potential of glycosylated carbon nanotubes in the delivery of radionuclides for imaging and radiotherapeutic purposes. Single-walled carbon nanotubes were filled with a radionuclide and after proper functionalisation decorated with different sugars to deliver them to different targets.

Finally, we dedicated one last section to the design of functionalised fullerenes as support for the solid-phase synthesis of oligosaccharides. Fullerene derivatives are able to perform homogenous solution phase synthesis with all the advantages of a solid-phase work-up through precipitation.

In this thesis we touched several aspects of carbohydrate synthesis, manipulation and applications, to provide new tools which will help elucidate the role of carbohydrates in biological systems.

Acknowledgements

In this very long journey I have accumulated many people to thank.

First, I would like to thank Professor Ben Davis, the ITN FP7 of the Marie Curie Actions, and Merton College for making this DPhil possible. I would also like to thank all the people I had the privilege to collaborate with (and learn from) such as Professor D. Anthony and Professor J. Schneider.

I am very grateful to the Department of Chemistry and all its staff for the tireless help and support towards my goals, visions and achievements both on practical matters as well as encouragement. In particular to C. Blackwell, the analytical facility staff, and my College Advisor Professor T. Softley.

My very deep gratitude to A. Bullard, T. Parsons, M. Sharma, B. Bhushan, C. Serpell, S. Sattin, R. Saliba, B. Bower, M. Bilyard, S. Gunnoo, K. Wals, J. Willwacher and Mumu, the army of proof-readers that helped me editing this thesis with all my mistakes. It has been a unique opportunity and a pleasure to spend some time with my examiners, Professor G. Fleet and Professor P. Crocker, discussing my science. Their profound knowledge of science and outstanding kindness is without parallel.

I would also like to thank the BGD group members who have been looking after me through the years and shared with me my highs, my lows, science and unforgettable moments of diversion. Within the group I had the chance to meet incredible people and form deep friendships. I would like to thank R. Rutte, for her infinite love and efforts in preserving my mental sanity; B. Bhushan, for all the adventures we embarked on together and his desperate attempts to fight my lack of self-preservation; T. Parsons, for his impeccable composure in those moments

when I went completely berserk, and for rescuing me (and my reactions) in a number of occasions. You all set a very high bar.

Pursuing my DPhil in Oxford has been an immense privilege, but furthermore a pleasure, thanks to a number of friends I have been fortunate to meet through rowing, volleyball, 11 Norham Gardens, and a number of activities in which I took part in the past few years. I will never forget the adventure that brought me to the roof-top of Africa and with it the people that transformed climbing the Kilimanjaro from a dream to reality. I would like to thank the Oxford Volleyball Club, and in particular my team, for the love and fun we have shared through our mutual passion. In the past years playing together I have always received from you more than I could give back. A thank you also to the Oxford Academicals Rowing Club, in which I finally found a squad to nourish my passion for rowing. In the past few years the way I look at rivers changed forever: on the water, I have experienced true beauty, brutal hardness, and the most complete happiness while sliding quietly like a leaf on the water in a single scull.

Finally, I would like to thank my family for their support over the past four years - and almost 26 years before that. There are no words to describe how much you have done for me, and how I am utterly incapable of expressing all my gratitude for that. For you I have the most unconditional love and the never fading hope that I will be able, one day, to show you how much more than this DPhil you made possible. I will never forget that a strong character is full of scars.

May my gratitude to all the people mentioned in this document become eternal, deposited in the Bodleian Library for years to come, along with my gratitude to all the people I could not include, who will live forever in my memories.

Authorship Declaration

I declare that this thesis represents my own work, except where stated below. It has never been submitted for any other degree.

FuT was expressed by Dr. Regis Saliba in the Benjamin Davis Laboratory, Chemistry Research Laboratory, University of Oxford.

ccHSQC spectra have been been acquired by Prof. Tim Claridge at the NMR facilities of the Department of Chemistry, University of Oxford

TTMA was synthesised by Peter Crane in the Benjamin Davis Laboratory, Chemistry Research Laboratory, University of Oxford.

SPR data were recorder by Dr. Torben Schiffner in the Quentin Settentau Laboratory, Sir William Dunn School of Pathology, University of Oxford

EM and TGA analyses of the carbon nanotubes were performed by Dr. Gerard Tobias at the Institut de Ciència de Materials de Barcelona.

Sections of tissues were embedded in wax, cut and mounted on glass slides by Richard Stillion at the histology facility of Sir William Dunn School of Pathology, University of Oxford.

Abbreviations

The following abbreviations have been used in this thesis:

Amino acids:

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Asparagine
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

Sugars:

Fuc	Fucose
Gal	Galactose
GalNAc	N-acetylgalactosamine
Glc	Glucose
GlcA	Glucuronic acid
GlcNAc	N-acetylglucosamine
IdoA	Iduronic acid
Lac	Lactose
LacNAc	N-acetyllactosamine
Le ^a	Lewis a
Le ^b	Lewis b
Le ^x	Lewis x
Le ^y	Lewis y
Man	Mannose
ManNAc	N-acetylmannosamine
Neu	N-acetylneuraminic acid or sialic acid
sLacNAc	sialylated LacNAc
sLe ^x	sialylated Le ^x
Xyl	Xylose

General:

2D	two-dimensional
3D	three-dimensional

Å	Angstrom
Ac	Acetyl
Ac ₂ O	Acetic anhydride
AcCl	Acetyl chloride
ACN	Acetonitrile
AcOH	Acetic acid
AgOTf	Silver triflate
aq	Aqueous
Ar	Aromatic
ASGP	Asialoglycoprotein
ATR	Attenuated total reflectance
BCA	Bicinchoninic acid
Bn	Benzyl
BSA	Bovine Serum Albumin
Bz	Benzoyl
°C	Degrees Celsius
Calc.	Calculated
CAN	Ceric Ammonium Nitrate
ccHSQC	Carbon-coupled HSQC
cm	Centimetre
CNTs	Carbon nanotubes
CRD	Carbohydrate Recognition Domain
CT	Computed Tomography
CTP	Cytidine triphosphate
CuAAC	Copper catalysed azide-alkyne cycloaddition

CVD	Carbon Vapour Deposition
δ	Chemical shift
Da	Dalton
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DC-SIGN	Dendritic Cell-Specific Intercellular Grabbing Non-integrin
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
DHP	Dihydropyran
DIPEA	N,N-Di-isopropylethylamine
DMAD	Dimethyl acetylenedicarboxylate
DMAP	4-(Dimethylamino)pyridine
DMF	Dimethylformamide
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
DTBMP	2,6-Di-tert-butyl-4-methylpyridine
DTBS	Di- <i>tert</i> -butyl-silyl
DTT	Dithiothreitol
E. coli	Escherichia coli
EDAC	<i>N</i> -Ethyl- <i>N'</i> -(3-dimethylaminopropyl)carbodiimide
EDX	Energy-Dispersive X-ray
e.g.	<i>exempli gratia</i>
EM	Electron Microscopy
ECM	Extracellular matrix
eq.	Equivalents
ER	Endoplasmic Reticulum

ESI	Electrospray ionisation
Et	Ethyl
et al.	<i>et alia</i>
FGD	Fluorodeoxyglucose
FI	Field Ionisation
Fig.	Figure
Fmoc	Fluorenylmethyloxycarbonyl
FuT	Fucosyltransferase
g	grams
GDP	Guanosine 5'-diphosphate
GNP	glyconanoparticle
GPI	Glycosylphosphatidylinositol
h	hours
HAADF	High-Angle Annular Dark-Field
HATU	1-[Bis(dimethylamino)methylene]-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i>] pyridinium 3-oxid hexafluorophosphate
HOBt	1-Hydroxybenzotriazole
Hpg	Homopropargylglycine
HPLC	High-Performance Liquid Chromatography
HRMS	High-Resolution Mass Spectrometry
Hz	Hertz
<i>i</i>	iso
ID	Injected Dose
IL	Interleukin
IME	2-imino-methoxyethyl

IPA	Isopropanol
IR	Infrared
<i>J</i>	Coupling constant
k	Kilo
K222	Kriptofix® 222
K _d	Dissociation constant
LC-MS	Liquid-Chromatography Mass Spectrometry
LevOH	Levulinic acid
LG	Leaving Group
LHMDS	Lithium bis(trimethylsilyl)amide
μ	Micro
M	Molar or molecular mass ion or mega
Man-R	Mannose receptor
Me	Methyl
mg	Milligrams
min	Minutes
μL	Microlitres
mL	Millilitres
MLs	Magnetoliposomes
mm	Millimetres
mM	Millimolar
mmol	Millimoles
mol	Moles
MRI	Magnetic Resonance Imaging
MS	Molecular Sieves or Mass Spectrometry

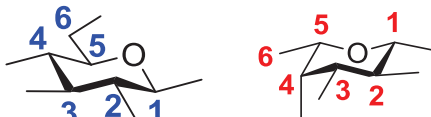
MW	Microwaves
MWNTs	Multi-walled carbon nanotubes
m/z	Mass/charge ratio
NaOMe	Sodium methoxide
NEt ₃	triethylamine
NHS	N-hydroxysuccinimide
NmCSS	<i>Neisseria meningitidis</i> CMP-sialic acid synthetase
NMR	Nuclear Magnetic Resonance
NP	Nanoparticles
Ns	Nosyl
NTs	Nanotubes
ODCB	Orthodichlorobenzene
OS	Oligosaccharides
PCC	Pyridinium chlorochromate
PE	Petroleum ether
PET	Positron Emission Tomography
PFA	Paraformaldehyde
PG	Protecting Group
Ph	Phenyl
pK _a	Acid dissociation constant
PMB	<i>para</i> -methoxybenzyl
PmST	<i>Pasteurella multocida</i> sialyltransferase
ppm	Parts per million
quant.	Quantitative
QDs	Quantum Dots

R	Generic organic group
R _f	Retention factor
RNA	Ribonucleic acid
r.t.	Room temperature
SCM	S-cyanomethyl
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	Scanning Electron Microscopy
SFB	N-succinimidyl-4-fluoro-benzoate
SMCs	Smooth muscle cells
SPECT	Single-Photon Emission Computed Tomography
SPOS	Solid-Phase Oligosaccharide Synthesis
SPPS	Solid-Phase Peptide Synthesis
SPR	Surface Plasmon Resonance
STEM	Scanning Tunnelling Electron Microscopy
SWNTs	Single-walled carbon nanotubes
<i>t</i>	<i>tert</i>
t _{1/2}	half-life
TBAB	Tetrabutylammonium bromide
TBAF	Tetrabutylammonium fluoride
TBAI	Tetrabutylammonium iodide
TBDMS	<i>tert</i> -butyldimethylsilyl
TBDPS	<i>tert</i> -butyldiphenylsilyl
TCA	Trichloroacetimidate
TEM	Transmission Electron Microscopy
Tf	Trifluoromethanesulfonyl (triflyl)

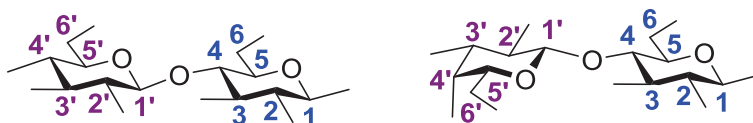
TFA	Trifluoroacetic acid
TGA	Thermogravimetric analysis
THF	Tetrahydrofurane
TLC	Thin Layer Chromatography
TMS	Trimethylsilyl
TPAH	Tetrapropylammonium hydroxide
TTMA	(tris(1-((O-ethyl)carboxymethyl)-1-H-1,2,3-triazol-4-ylmethyl) amine)
UV	Ultraviolet
WGA	Wheat Germ Agglutinin

Explanatory Note

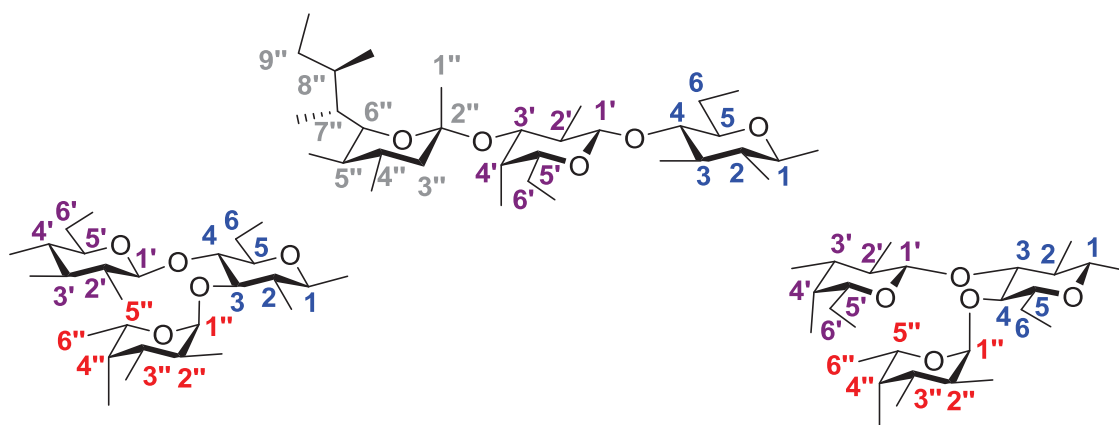
Monosaccharides



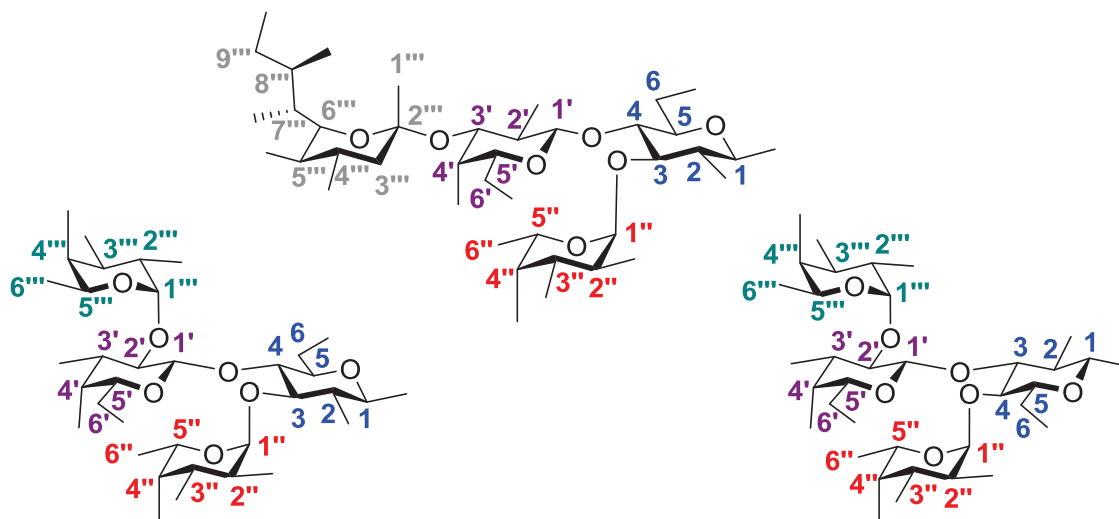
Disaccharides



Trisaccharides



Tetrasaccharides



General Considerations

Proton nuclear resonance magnetic (^1H NMR) spectra, ^{13}C NMR, correlation spectroscopy (COSY), heteronuclear single-quantum correlation spectroscopy (HSQC), and carbon-coupled HSQC (ccHSQC) were recorded on Bruker AV500, Bruker AV400, Bruker AV250 and Bruker AV200 (500, 400, 250, 200 MHz) spectrometers. All chemical shifts are quoted on the δ scale in ppm, using residual solvents as the internal standard (^1H NMR: CDCl_3 = 7.26, DMSO-d_6 = 2.50, MeOD = 3.31, CD_3CN = 1.94, D_2O = 4.79, toluene- d_8 = 2.08, 6.07, 7.01, 7.09, $(\text{CD}_3)_2\text{CO}$ = 2.05; ^{13}C NMR: CDCl_3 = 77.16, DMSO-d_6 = 39.52, MeOD = 49.00, CD_3CN = 1.32, toluene- d_8 = 2.43, 125.13, 127.96, 128.87, 137.48, $(\text{CD}_3)_2\text{CO}$ = 29.84, 206.26). Coupling constants (J) are reported in Hz with the following splitting abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublets, dt = doublet of triplets, qd = quartet of doublets, ddd = doublet of doublet of doublets, br = broad. All chemical shifts were assigned using COSY and HSQC spectra.

Low resolution mass spectra (MS) and high resolution MS (HRMS) were recorded on a Waters Micromass LCT Premier TOF spectrometer using electrospray ionisation (ESI) or field ionisation (FI). Liquid Chromatography-Mass Spectrometry (LC-MS) was performed on a Waters LCT Premier XE (ESI-TOF-MS) coupled to a Waters 1525 Micro HPLC pump 2777 autosampler, using a Merk Chromolith RP-18 2 x 5 mm guard column. Water containing 0.1% formic acid by volume (solvent A) and acetonitrile (solvent B), were used as the mobile phase at a flow rate of 0.4 mL/min. The gradient was programmed as follows:

- 0 min 95% A 5% B 0.4 mL/min

-	1 min	95% A 5% B	0.4 mL/min
-	5 min	0% A 100% B	0.4 mL/min
-	8 min	0% A 100% B	1 mL/min
-	8.1 min	95% A 5% B	1 mL/min

The electrospray source was operated with a capillary voltage of 3 kV and a cone voltage of 100 V. Nitrogen was used as the nebulizer and desolvation gas, at a total flow of 300 L/hr. Spectra were calibrated using a calibration curve, constructed from a minimum of 17 matched peaks from the multiply charged ion series of myoglobin, obtained at a cone voltage of 100 V. Total mass spectra were obtained from the ion series using the MaxEnt algorithm, pre-installed on MassLynx software (v. 4.0 from Waters) according to the manufacturer's instructions. Myoglobin was purchased from Sigma-Aldrich. SDS-PAGE gels were run using pre-cast gels purchased from Invitrogen (NuPAGE 4-12% Bis-Tris gel), in MES buffer, 200 V for 35 min.

Optimal rotations were measured on a Perkin-Elmer 241 polarimetry with a with a path length of 1 dm and are reported in units of 10^{-1} deg cm² g⁻¹. Concentrations are given in g/100 mL. Infrared (IR) spectra were recorder on a Bruker Tensor 27 Fourier Transform spectrophotometer with a diamond ATR module. Absorption maxima are reported in wavenumbers (cm⁻¹). Thin layer chromatography (TLC) was carried out using Merk aluminium backed sheets coated with 60 F254 silica gel or alumina. Visualisation of the silica plates was achieved using a UV lamp (λ_{max} = 254 nm), and/or ammonium molybdate (5% in 2 M H₂SO₄), ninhydrin (1.5 g ninhydrin, 5 mL acetic acid, 500 mL 95% ethanol) or sulfuric acid (45 mL methanol, 45 mL water, 3 mL sulfuric acid) solutions. Flash column

chromatography was carried out using Geduran Si 60 (40-63 μm) Merk. Mobile phases are reported as ratio volume:volume.

All solvents were used as supplied (analytical or HPLC grade), without prior purification. Deionised water was used for chemical reactions and protein manipulations. Reagents were purchased from Sigma-Aldrich and used as supplied. PE refers to the fraction of petroleum ether boiling in the range of 40-60 $^{\circ}\text{C}$. All reactions requiring anhydrous conditions we performed inflame-dried flasks, with activated molecular sieves, dried solvents and argon atmosphere.

Brine refers to a saturated aqueous solution of sodium chloride. Anhydrous magnesium sulphate (MgSO_4) was used as drying agent after reaction work-up. All compounds already described in literature are reported in the supporting information of each chapter with a superscript reference on the name of the compound. The reference indicates the paper where that compound has been previously described and where the comparative data can be found.

Animals were purchased from Charles River (UK); animal experiments were performed by Sonia De Munari (PIL: IC7F41F8D) under the project license 30/3076 with UK Home Office and local ethical approval.

Room temperature was defined as 21 $^{\circ}\text{C}$.

Lectin binding of glycoconjugates was tested on a biacore 3000 (BIAcore, Uppsala, Sweden) surface plasmon resonance (SPR) biosensor at 25 $^{\circ}\text{C}$ using HBS-EP buffer (GE Healthcare UK Ltd, Little Chalfont, United Kingdom).

All biological sterile manipulations were undertaken in a HERA safe KSP12 laminar flow hood (Thermo Scientific).

Microscopy experiments were performed on a Leica Microsystems SP5 Inverted Confocal Microscope. All experiments were performed with the pinhole set at 1

Airy diameter. All images were captured at 512 x 512 pixels at 400 Hz or 1400 Hz capture rates. All experiments were performed with the 20 x dry objective with further magnification being achieved with the optical zoom. Images were processed with the Leica LAS software or ImageJ. Stained slides sections were either recorded with a Leica Microsystems SP5 Inverted Confocal Microscope (no DAPI laser) or with a Leica CM5000B.

Table of Contents

Abstract.....	i
Acknowledgements.....	iii
Authorship Declaration.....	v
Abbreviations.....	vii
Explanatory Note.....	xvii
General considerations.....	xix
Table of Contents.....	xxiii

Chapter 1: Introduction.....1

1.1	The Glycocode.....	1
1.1.1	Sugars in nature.....	1
1.1.2	The role of sugars in biology.....	4
1.1.3	The interest of the community in sugars.....	8
1.2	Multivalency and the cluster glycoside effect.....	11
1.3	Glycosylated macromolecular structures.....	13
1.3.1	Macromolecules in biology.....	13
1.3.2	Synthetic macromolecular systems.....	14
1.3.3	Glycosylation of macromolecular structures.....	15
1.4	Aim.....	16
1.5	References.....	18

Chapter 2: Carbohydrate synthesis.....25

2.1	Introduction.....	25

2.1.1	Structure and nomenclature of carbohydrates.....	25
2.1.2	Synthesis of carbohydrates.....	29
2.1.3	Characterisation of carbohydrates.....	31
2.2	Glycomap sugar library.....	33
2.3	Carbohydrate synthesis.....	36
2.3.1	Monosaccharides.....	36
2.3.2	Disaccharides.....	41
2.3.3	Trisaccharides.....	42
2.3.4	Tetrasaccharides.....	48
2.4	2-Imino-2-methoxyethyl (IME) reagents.....	52
2.5	Conclusions.....	54
2.6	References.....	55
2.7	Supporting Information.....	58

Chapter 3: A triply divergent reagent for glycoprotein synthesis.....167

3.1	Introduction.....	167
3.1.1	Synthesis of glycoproteins by a triply divergent reagent.....	167
3.1.2	Site-specific protein modification.....	169
3.2	Protein scaffold.....	170
3.2.1	Homopropargylglycine (Hpg).....	170
3.2.2	Np276-M61Hpg.....	171
3.2.3	Q β -M16Hpg.....	172
3.3	A triply divergent reagent.....	173
3.3.1	From CN to IME.....	173

3.3.2	From CN to NCS.....	174
3.3.3	From CN to N ₃	175
3.4	Protein glycosylation.....	176
3.4.1	Protein glycosylation via IME.....	176
3.4.2	Protein glycosylation via NCS.....	177
3.4.3	Protein glycosylation via N ₃	177
3.5	Binding experiments.....	178
3.6	Conclusions.....	180
3.7	References.....	181
3.8	Supporting Information.....	184

Chapter 4: Iron nanoparticles as multimeric glycosylated scaffolds.....209

4.1	Introduction.....	209
4.1.1	Molecular imaging techniques: PET and MRI.....	209
4.1.2	Glyconanoparticles as multivalent scaffolds.....	211
4.1.3	Dual MRI-PET probes.....	212
4.2	Design and synthesis of glyconanoparticles.....	214
4.3	Characterisation of glyconanoparticles.....	215
4.4	Glyconanoparticle library.....	218
4.4.1	Sugar equivalents and coverage.....	218
4.4.2	Glyconanoparticle library.....	220
4.5	Synthesis of [¹⁸ F]-GNPs.....	221
4.5.1	Synthesis of [¹⁸ F]-SFB.....	221
4.5.2	Synthesis of [¹⁸ F]-GNPs.....	224

4.6	Conclusions.....	226
4.7	References.....	227
4.8	Supporting Information.....	231

Chapter 5: Nanocapsules for targeted delivery of radioactivity.....255

5.1	Introduction.....	255
5.1.1	Radiotherapy.....	255
5.1.2	Carbon nanostructures.....	256
5.2	Synthesis of glyconanotubes.....	260
5.2.1	Purification and filling of single-walled carbon nanotubes.....	260
5.2.2	Functionalisation of NaI@SWNTs.....	263
5.2.3	Design of a new linker.....	272
5.3	Characterisation of glyconanotubes by electron microscopy.....	276
5.4	Synthesis of Na ¹²⁵ I@SWNTs.....	279
5.5	Conclusions.....	281
5.6	References.....	282
5.7	Supporting Information.....	288

Chapter 6: Biological experiments.....335

6.1	Introduction.....	335
6.1.1	Carbohydrate-binding structures.....	335
6.1.2	Carbohydrate-binding specificity.....	340
6.2	Glycomap in vivo experiments.....	343

6.2.1	<i>In vivo</i> experimental set up.....	343
6.2.2	GNPs distribution and statistical analysis.....	347
6.2.3	Histological studies.....	360
6.3	<i>In vivo</i> testing of ^{18}F -GNPs.....	361
6.4	<i>In vivo</i> testing of glycosylated Na^{125}I @SWNTs.....	362
6.5	Conclusions.....	364
6.6	References.....	366
6.7	Supporting Information.....	374

Chapter 7: Fullerenes as support for oligosaccharide synthesis.....449

7.1	Introduction.....	449
7.1.1	Solid-phase synthesis of oligosaccharides.....	449
7.1.2	Fullerenes as support for SPOS.....	451
7.2	Synthesis of fullerene tags.....	453
7.2.1	First approach: functionalisation with a Grignard reagent.....	453
7.2.2	Second approach: zwitterion mediated functionalisation.....	454
7.2.3	Third approach: functionalisation by ethynylation.....	457
7.3	Conclusions.....	460
7.4	References.....	461
7.5	Supporting Information.....	462

Summary.....487

Appendix

INTRODUCTION

1.1 The Glycocode

1.1.1 Sugars in nature

Glycoscience is the discipline devoted to the study of carbohydrates, from their chemistry to their biological significance and applications.^[1] The sum of the interactions of sugars in biological systems forms the glycocode, which describes the complex pool of information that carbohydrates are able to express.^[2]

In nature, glycosylation is the process developed to decorate different structures (such as proteins and lipids) with glycans, to impart to them new functions, structural and physico-chemical properties.^[3] This process is dynamic and begins when a protein is targeted to the endoplasmic reticulum (ER), where dedicated enzymes attach the first set of glycans. Other glycosylation enzymes then either add or trim these sugar moieties, building a platform for the addition of new carbohydrates to a different part of the growing glycosylated structure. This back-and-forth process of glycosylation continues into the Golgi apparatus, which sorts the new proteins, distributing them to their final destinations in the cell (Fig. 1.1).

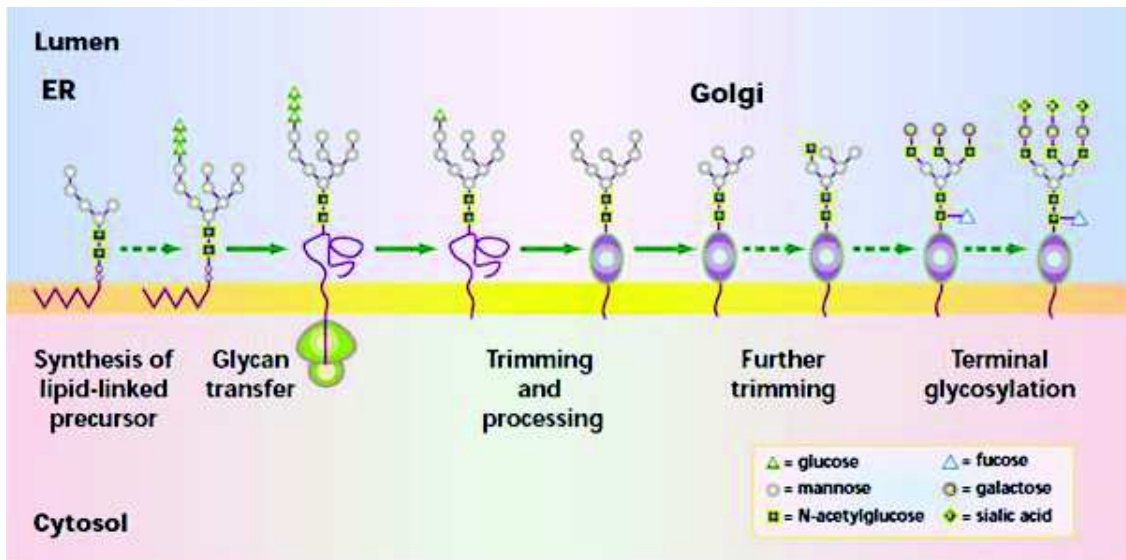


Figure 1.1 Protein glycosylation pathway.^[4] In the ER dedicated enzymes attach the first glycans. Into the Golgi other enzymes complete the glycosylation process.

In human cells, fucose (Fuc) and sialic acid (Neu) sit in a privileged position in that they often serve as a final cap for the tips of the carbohydrate branches.^[5]

Sugars are conjugated to proteins in three ways:

- N-linked glycans (onto Asn residues);
- O-linked glycans (onto Ser and Thr residues);
- Glycosylphosphatidylinositol (GPI) anchors.

N-linked glycans, containing the pentasaccharide Man- α -1,6-Man- α -1,3-Man- β -1,4-GlcNAc- β -1,4-GlcNAc as a common core, can be classified into four main groups: oligomannose (or high mannose), complex, hybrid, and poly-N-acetyllactosamine (poly-LacNAc).^[6] Oligomannose-type glycans contain only α -mannosyl residues attached to the trimannosyl core. Complex-type glycans contain the trimannosyl core and have branches with N-acetyl glucosamine (GlcNAc) residues. The hybrid-type N-glycans have the features of both complex-type and high mannose-type glycans. The poly-LacNAc N-glycans contain

repeating units of (Gal- β -1, 4-GlcNAc- β -1,3 or LacNAc) attached to the core.^[7] For GPI anchored proteins, the linkage is via the phosphatidylinositol on proteins located on the surface of cells. The resulting extent and position of the glycosylation depends on the conformation of the protein, conformation of the glycosylation site and enzymes involved (specific for cell types and tissues). Identical protein backbones bearing different glycosylation patterns are defined as glycoforms. Producing well-defined glycosylated protein structures via biochemical approaches is often a challenge as it typically leads to a mixture of inseparable glycoforms.

In order to achieve a better understanding of the abundance and distribution of sugars in nature, bioinformatics analyses have scanned the mammalian oligosaccharides present in our databank. From all the structures known, there is apparently no difference between the human and the mammalian set of oligosaccharide structures.^[8] The analysis showed that only 36 building blocks are necessary to synthesize about 75% of the mammalian glycome,^[9] and 224 to construct all theoretically possible mammalian oligosaccharides by linear chemical synthesis. The average oligosaccharide is composed of about 8 sugar units, whereas in the whole database structures vary from 1 to 37 monosaccharide units in size, with most mammalian structures (95%) having fewer than eight residues. As many as 88% of all glycosidic linkages found can be built from only 20 monosaccharide building blocks,^[9] the most frequently occurring connections being 4-linked β -GlcNAc, capping α -Neu, capping α -Fuc, capping β -galactose (β -Gal), 2-linked α -mannose (α -Man), and 3-linked β -Gal. Further comparative analysis of the mammalian glycome^[10,11] revealed that the most abundant monosaccharides are β -D-GlcNAc (32%), β -D-Gal (25%), α -D-

Man (19%), α -Neu, α -L-Fuc and β -D-Man. Altogether they cover more than 75% of all monosaccharide units found. Astonishingly D-Glc, the most abundant monosaccharide in nature (making up cellulose, starch, and glycogen), plays only a minor role within mammalian O- and N-glycans. Three different monosaccharides dominate the glycans' non-reducing terminus, the site most often recognized by carbohydrate-binding proteins: D-Gal, Neu, and L-Fuc each cap 25% of oligosaccharides. D-Man and D-GlcNAc each terminate 8% of mammalian oligosaccharides. Finally, amongst all the possible modifications of the monosaccharide, N-acetylation is the most abundant, followed by O-acetylation.

1.1.2 The role of sugars in biology

Almost all cell surfaces and secreted proteins are glycosylated. The structures of their glycans are implicated in protein folding, cell signalling, fertilisation, embryogenesis, neural development, hormone activity, cell proliferation, pathogen recognition, inflammation, immune response, the development of autoimmune disease and cancer.^[7,12,13] Many of these roles are covered by sialic acid and as a consequence, this sugar has attracted the attention of researchers. Sialic acid engineering endeavours to mimic nature, which uses over 50 different forms of Neu to modulate the structure and function of sialylated proteins and lipids.^[14] The development of N-acetylmannosamine (ManNAc) analogues (the first committed intermediate in sialic acid synthesis) for general use in sialic acid engineering applications has been previously explored by Jones et al.^[15] to investigate the sialic acid biosynthetic pathway. Moreover, by using metabolic

substrate-based methods, researchers can design completely new structures and impart completely novel functions to biological systems. Similarly to Neu analogues (and for the same reasons), Fuc analogues have also been developed.^[16]

To understand the biological relevance of monosaccharides, an immense amount of work has been done over the last few decades,^[17] developing new glycoproteins, nanosystems and constructs to gain a greater understanding of the sugar code.^[18,19] The ability to visualise and isolate cellular glycans and their interacting partners would help to deconvolute the complexity and micro-heterogeneity that make it difficult to study their biological function. Towards this goal, labelled glycoproteins or artificially produced neoglycoproteins have been used to detect endogenous carbohydrate-binding proteins.^[20] In 1992, a study compared the *in vivo* binding of a set of fluorescently-labelled neoglycoproteins^[21] (displaying β -Gal, Fuc, GalNAc, GlcNAc, Man, Xyl) after intravenous injection by collecting samples of tissue sections (brain, lungs, liver, spleen and kidney), to assess the validity of neoglycoprotein histochemistry. One approach to follow monosaccharide distribution, and to do *in vivo* kinetic analysis, used chemically prepared neoglycoproteins in which Gal, Man and Fuc (amongst others) were conjugated onto the protein surface by use of chemical methods.^[22] Their kinetic and organ-specific accumulation have been extensively studied after intravenous injection and analysed by conventional dissection studies, as well as non-invasive methods such as SPECT (single-photon emission computed tomography), PET (positron emission tomography), or fluorescence imaging. These studies revealed the affinity of Gal-modified proteins for asialoglycoprotein receptors on liver parenchymal cells. Man, on the other hand, was found to tune

the neoglycoproteins' affinity to non-parenchymal cells, (e.g. Kupffer cells) or to the endothelial cells that express the mannose receptors. In 2007 Hsu and co-workers^[23] incorporated alkyne groups into cellular glycans by appending Fuc and ManNAc alkyne derivatives. The remodeled glycans, which contained these bioorthogonal chemical handles, were then chemoselectively labeled with a complementary reactive probe for further manipulation, visualization and isolation. A Fuc-modified neoglycoprotein labelled by ¹¹¹In was reported to accumulate in the liver, especially in the Kupffer cells.^[24,25] By SPECT, Man modified proteins were found in non-parenchymal cells whereas Gal congeners were captured by hepatocytes, lymph nodes, liver and spleen.^[26] Kobayashi and co-workers^[27] compared the effects of monosaccharides on protein kinetics by non-invasive fluorescence imaging; human serum albumin protein conjugated with Glc, Gal, Man, and Fuc has been imaged in nude mice using near-infrared fluorescence detection. Glc and Gal modified albumins were found to be captured by hepatocytes, while the Man and Fuc conjugates were catabolized in macrophages. *In vivo* testing of iodine labelled protein–liposome conjugates in Ehrlich solid-tumor-bearing mice^[28] has also been used to study sLe^x biodistribution (Fig. 1.2).

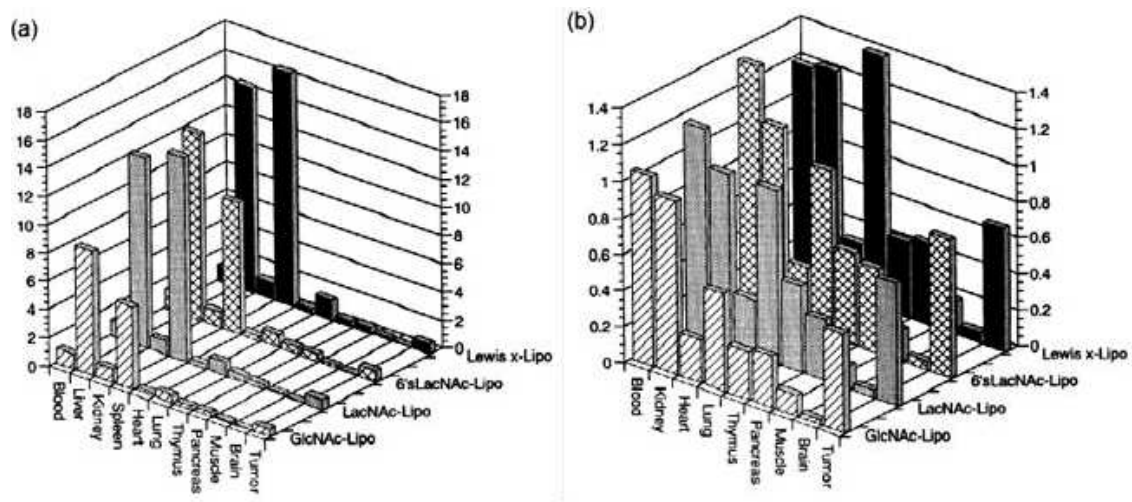


Figure 1.2 sLex^x *in vivo* biodistribution studies.

A very important feature of carbohydrates, which led to the discovery of sugar-based drugs, is their involvement in the immune and inflammatory response.^[29] White blood cells (leukocytes) circulate throughout the blood vasculature, patrolling the body for sites of tissue damage or microbial infection. When such a site is found, the leukocytes exit the blood stream and migrate to the surrounding tissue to perform their protective functions. The sequence of events leading to leukocyte recruitment, the inflammatory cascade,^[30] begins with the interaction of circulating leukocytes with endothelial cells on the blood vessels. Leukocytes roll along the endothelial layer propelled by the blood flow, adhere to it and cross it to migrate into the tissue. Excessive and uncontrolled immune response results in inflammatory diseases and autoimmune diseases. When it was discovered that the initial rolling event is mediated by carbohydrate binding receptors, sugar chemists opened up a new era in anti-inflammatory drug design.

1.1.3 The interest of the community in sugars

The structural biology of oligosaccharides has undergone a rapid expansion in the first few years of the 1990s, from interesting only a small number of researchers to engaging great areas of biological chemistry and medicine. Despite this growing feeling of the importance of carbohydrates in the scientific community, there are few carbohydrate-based drugs in the market^[31] due to the high polarity of sugars leading to poor pharmacokinetic properties.^[32] As drugs, sugars are poor candidates as they have high hydrophilicity, and as a result, poor membrane permeability and short serum half-life. Nevertheless, scientists and pharmaceutical companies invested steadily in the development of carbohydrate-based vaccines.^[33] Glycovaccines have been used to induce immunity since the 19th century, when Louis Pasteur first pioneered their development. Sugar-based vaccines are still mainly obtained from extraction of the polysaccharidic moiety from an organism, which is then conjugated to a carrier at a later stage (e.g. anti-meningococcal vaccine and anti-*Haemophilus influenzae* type B vaccines). Fully synthetic glycovaccines are not marketed to date but research is unrelentingly moving in this direction.^[34-36] Some types of glycovaccines have not experienced much scientific impetus: this is the case for anti-fungal and parasite vaccines, which suffered a loss of interest after the discovery of the malaria vaccine.^[33] Other glycovaccines have instead conquered the heart of vaccine design. Cancer cells present aberrant glycosylation either in structure or copy number, and cancer vaccines have become the new focus for development over the last few years.^[37-50] Sadly no cancer vaccines are available yet, though some have finally reached clinical trials.^[48] The use of carbohydrates as vaccines has also been

hampered by their poor immunogenicity. In designing appropriate formulations, the introduction of adjuvants plays a key role in resuscitating the interest in these carbohydrate-based drugs.^[51]

Another application of sugars is their use as drug delivery systems in 'glycotargeting'. The concept of targeted drug delivery was proposed by Paul Ehrlich more than 100 years ago.^[52] In glycotargeting, carbohydrate moieties are to be used as delivering agents of drugs or imaging probes, and have been the subject of fairly intensive exploration since the 1990s.^[53-55] Amongst the commonly used carbohydrate targeting ligands^[26] β -linked Gal or GalNAc, binding to the hepatic asialoglycoprotein (ASGP) receptor, have been studied in the greatest detail.^[56,57] Other receptors have been isolated and their interacting partners have been identified (Table 1.1); this is the case for the mannose receptor (Man-R), highly expressed on antigen-presenting dendritic cells (DC-SIGN) as well as on macrophages and hepatic endothelial cells, which mediates crucial biological processes such as endocytosis and the processing of glycoprotein antigens that expose D-Man and L-Fuc residues.^[58]

Lectin	Saccharide Specificity	Core motif	Location	Direct Interaction to Viral Components
C-type lectin	Variable	C-type sequence motif	Extracellular, Cell membrane	Yes
Galectin	β -Galactosides	S-type sequence motif	Extracellular, Cytoplasm	Yes
Calnexin	Glc ₁ Man ₉	Calnexin sequence motif	ER	Yes
P-type lectin	Mannose-6-P, others	P-type sequence motif	Cell membrane, Endosome	Yes
L-type lectin	Variable	L-type sequence motif	ER, ERGIC, Golgi	Yes
I-type lectin	Sialic acid, variable	Ig-like domains	Cell membrane	No
M-type lectin	Mans	M-type sequence motif	ER	No
F-type lectin	L-fucose	F-type sequence motif	Extracellular	No
R-type lectin	Variable	R-type sequence motif	Extracellular, Cell membrane	No
F-box lectin	GlcNAc ₂	F-box sequence motif	Cytoplasm	No
Chitinase-like lectin	Chito-oligosaccharides	TIM (Triose-phosphate isomerase) barrel-like structure	Extracellular	No
Intelectin	Gal, galactofuranose, pentoses	Intelectin sequence motif	Extracellular, Cell membrane	No

Table 1.1 Classification of mammalian lectins, location and sugar-binding specificity.^[59]

In 2010 Ma et al.^[60] introduced Lac moieties onto nanoscale micelles composed of biodegradable amphiphilic block copolymers, encapsulating a fluorescent dye. They took the merits of both polymer–drug conjugates and ligand-receptor targeting strategies to achieve the display of multiple functionalities, water solubility, prolonged circulation, biodegradability, liver cell targeting, rapid cell uptake, and fluorescent probing. Because lactose is reported to have a specific interaction with the ASGP receptors over-expressed on hepatocyte membranes,^[61-63] these micelles were expected to have liver-targeting ability. Mice were injected with Lac-micelles and Lac-free micelles as a control. Immediately after sacrifice, liver, kidneys, spleen, lung, and heart were excised from each mouse, homogenised and analysed, together with histological sections of each organs. Results showed that both Lac and Lac-free micelles had similar distribution, however comparatively, the liver of the mouse injected with Lac-micelles seemed to contain more micelles than that of the mouse injected with Lac-free-micelles. Homogenates of the excised organs showed the following distribution: liver >> kidneys and spleen > lungs > heart. On the other hand, the fluorescence intensity order was spleen and kidneys > liver > heart and lungs, not matching either what was observed in direct organ imaging nor in the histological studies. As always, comparing different techniques is essential to confirm the reliability of the data collected. Another example is the work of Benito-Alifonso and co-workers^[64] in which quantum dots (QDs) were decorated with Gal, Glc, Man and Lac to probe protein–carbohydrate interactions, and to identify carbohydrate receptors or ligands associated with intercellular recognition processes. Finally, Ketkar-Atre and co-workers^[65] tried to assess lactose functionalized magnetoliposomes (MLs) as a magnetic resonance imaging (MRI)

contrast agent to target the ASGP receptor *in vivo*. MRI signal intensity did not show significant differences between functionalized and non-functionalized particles, while iron-specific Prussian blue staining and transmission electron microscopy (TEM) analysis confirmed the uptake of lactose MLs mainly in hepatocytes. By contrast, non-functionalized anionic MLs were mainly taken up by Kupffer and sinusoidal cells.

What is clear is that regardless of the form in which sugars are entering the global scenario, the pressure to know more about carbohydrate properties and functions is surging.

1.2 Multivalency and the cluster glycoside effect

Multivalency plays a pivotal role in biological events concerning molecular recognition: protein-protein, protein-cell membrane, virus-cell and intercellular interactions.^[13,66-68] Complex theoretical models have been postulated to describe this effect,^[66,69,70] which can be explained kinetically.^[68] The binding rate of the first interaction with the ligand in a multivalent system is similar to the binding rate of the corresponding monovalent interaction, but the second interaction is faster since the ligand is placed in close proximity of the target by the system already bound. A similar effect is observed when there is an increased local concentration of the ligand or during an intramolecular interaction. This also means that if one interaction is lost it will be quickly re-established for the same reasons as outlined above.

Cooperativity is a different phenomenon, which arises from two or more interactions^[71] leading the system to deviate from the behaviour expected if the interactions are considered independently. This, in contrast to what happens

during a multivalent interaction, is due to the fact that interactions affect each other, as the first interaction modifies the binding site to facilitate the following one.

Most sugars bind weakly to receptors, with K_d values in the millimolar range.^[66,63] This is due to the fact that carbohydrate binding proteins usually have a shallow binding pocket for their ligands and therefore the resulting interaction is neither highly specific nor particularly strong. Lee and co-workers^[72] proposed that nature bypasses this limitation by employing multivalent carbohydrate ligands and coined the term 'cluster glycoside effect' to describe this phenomenon. Although this effect is largely accepted and profoundly described in the literature, there is still no agreement on its physical basis. The fact that carbohydrate-binding proteins have rather broad binding specificities for structurally related oligosaccharides (unlike other molecules like antibodies) implies that the sugar ligand for a given protein may cross-react with other proteins in the body. This is highly undesirable for therapeutic purposes because of the extensive side-effects that this would cause. Thus, for oligosaccharides to be considered in drug design, it is not enough to identify the sugar ligand of a single protein - rather, it would be more important to build a database of oligosaccharides that serve as ligands for proteins that occur in various locations in the body, to aid the design of specific analogues.

1.3 Glycosylated macromolecular structures

1.3.1 Macromolecules in biology

The term macromolecule was coined in the 1920s by the Nobel laureate Hermann Staudinger.^[73] At that time it had a different meaning than at present, as it was used to describe structures and had little to do with size.

Different disciplines give different meanings to macromolecules: while biology refers to macromolecules as those biopolymers present in living beings and essential for their biological function, in chemistry the term may refer to aggregates of two or more molecules held together by intermolecular forces (rather than covalent bonds) which do not readily dissociate. Macromolecules in biochemistry are biopolymers (nucleic acids, proteins and carbohydrates) and large non-polymeric molecules (such as lipids). Often the word 'polymer' is used synonymously with macromolecule, since macromolecules are commonly created by polymerization of smaller subunits (monomers). DNA, RNA, proteins and glycans are in fact repeating structures of related building blocks: nucleotides, amino acids and monosaccharides. Natural proteins are polymers of the 20 amino acids present in living organisms. Their expression takes place in the ribosome and it is controlled by stimuli activating the transcription and translation of genetic information. Synthetic proteins may carry unnatural amino acids and can be prepared by multistep synthesis or expressed by microorganisms. Glycans are made of linked units of monosaccharides, prepared in the cytoplasm and endoplasmic reticulum by dedicated enzymes.

Macromolecules often have unusual physical properties that do not characterise smaller molecules. One common property is their poor resistance to mechanical

stress; physical forces exerted on these molecules can overcome the strength of their covalent bonds and break the structure.^[74] Another macromolecular feature that does not apply to smaller molecules is their tendency to form colloids in water rather than proper solutions. This can be prevented by adding salts or particular ions to the preparation. Unfortunately, many proteins will denature if the concentration of the solute in the mixture is too high or too low, and crash out in hopeless precipitates. High concentrations of macromolecules in a solution can alter the rates and equilibrium constants of the reactions of other macromolecules, through an effect known as macromolecular crowding.^[75,76] This originates from macromolecules excluding significant quantities of other molecules from the solution, effectively increasing the concentrations of these molecules in areas occupied by the macromolecules.

1.3.2 Synthetic macromolecular systems

Synthetic macromolecules are a class of compounds created from the assembly of singular units into contiguous extensive structures. The two main platforms fitting this category are carbon-based structures and polymers. The first of these consist of graphene and other hyperconjugated supramolecular structures (fullerenes and nanotubes), which will be described more in detail in chapter 5. Polymer is instead a more general term indicating a construct of repeating monomeric units. The chemical structure of the units determines the identity of the polymer and, to some extent, its properties. As stated in the previous section, macromolecules do not exactly adopt all the physico-chemical properties of their constituent building blocks; due to their size, accessibility and interactions they

may present completely new ones. The properties of polymers depend also on chain length (the size of the molecule), side groups (responsible for intra- and intermolecular interactions), branching and cross-linking amongst chains. Polymers are defined as homopolymers if the units constituting the material are identical, or copolymers if they are different. Many synthetic polymers have been used for biomedical applications, in particular in organ and tissue regeneration.^[77] Supramolecular assembly of atoms or molecules led to a group of macromolecular structures more appropriately defined as supramolecular systems.^[78] Inspired by biological systems, this huge class of compounds generated the branch of science known as nanotechnology. In supramolecular structures, covalent bonds are still present in the constituent monomers, but the real protagonists of the assembly are weak intermolecular forces such as van der Waals, π - π stacking, hydrophobic forces, electrostatic effects, metal coordination and hydrogen bonding. Supramolecular structures can be entirely synthetic, such as metal and non-metal nanoparticles, nanotubes and nanorods,^[79] but also biologically-derived, such as DNA-based constructs.^[80]

1.3.3 Glycosylation of macromolecular structures

Glycosylation plays a dual structural and functional role in the stability, activity and interactions of macromolecular structures.^[81] Decorating structures with glycans can bring many advantages, from improving their solubility and biocompatibility^[82] to the addition of entirely new features, as a matter of targeting, circulation half-life and immunogenicity.^[83] To impart the desired properties to a structure through glycosylation a careful design has to be planned, considering

which carbohydrate moieties are the most appropriate and where to install them. Unfortunately even the most precise rational design of glycosylated structures is often insufficient for the challenges of *in vivo* testing. It is beyond our capacity to take into account all the possible complex entities they will encounter once inside an organism *a priori*, and often the optimisation goes through a process of trial and error. Even preliminary *in vitro* experiments are highly inadequate in this sense. Nevertheless, their importance to establish if the experimental design is based on solid grounds is incontestable.

1.4 Aim

The more knowledge we acquire on the structure and roles of carbohydrates, the more we become aware of the gaps we still need to fill to have the ability to detect, interpret, and correct (when necessary) biological processes. While our understanding of carbohydrate structures has shown a steady progression, our perception of their role in nature is still very limited. Part of the reason is that we find plenty of specific examples of carbohydrates and glycosylated structures involved in – again - specific functions, but very few general illustrations of their overall engagement. What we lack is an overview of key aspects that would collectively build a platform, the bigger picture, for us to explore in all its details at a later stage. To create this universal vision the aspects that need to be addressed are glycan structural diversity (glycoforms), distribution, level of expression, binding partners and function. A few of these topics have been covered quite thoroughly by researchers in the last few decades (as presented in this chapter), while others have been (unsurprisingly) lagging behind, probably because of the amount of effort, time and resources required. It is also important

to note how the attractiveness of the applicability of the findings often represents a key factor in dictating the direction that research efforts might take.

In this thesis our main aim was to challenge one very important feature of carbohydrates, the *in vivo* distribution of their natural targets. With a selected pool of saccharides in our hands, we tracked their distribution in mice, building what we called the 'Glycomap'. On the road to mapping these saccharide interactions, we explored two of their other important applications, as protein glycosylating agents and as targeting agents for drug delivery. After selecting a modest library of sugars, from mono to tetrasaccharides, we described their synthesis and application on different macromolecular systems. The synthesis of complex oligosaccharides was performed exploring both chemical and enzymatic glycosylation strategies. One sugar reagent has been modified in a divergent manner to exploit different protein glycosylation strategies on two different protein scaffolds. The whole library has been coupled to iron nanoparticles, to follow the sugar distribution *in vivo* and create the Glycomap. Sugars have also been used to decorate single-walled carbon nanotubes for the targeted delivery of radioactivity contained within their cavity. Finally, from the technological background developed in the previous chapters, we concluded with the investigation of C₆₀ fullerenes as support for the automated synthesis of oligosaccharides.

We hope our Glycomap will offer a solid, if modest, starting point for the determination of the distribution of sugar ligands and their biological partners *in vivo*. We envisage that these systems will offer the flexibility required in the future to isolate carbohydrate-binding structures, elucidate the role of sugars in biological systems and offer a platform for imaging and drug delivery applications.

1.5 References

- [1] G-J. Boons, Y. C. Lee, A. Suzuki, N. Taniguchi, A. G. J. Voragen, in *Comprehensive Glycoscience* (Ed.: H. Kamerling), Elsevier, **2011**.
- [2] J. Sebestik, M. Reinis, J. Jezek in *Biomedical Applications of Peptide-, Glyco- and Glycopeptide Dendrimers, and Analogous Dendrimeric Structures* (Ed.: H. Kamerling), Springer, **2012**, 93-98.
- [3] S. Defaus, P. Gupta, D. Andreu, and R. Gutiérrez-Gallego, *Analyst* **2014**, 139, 2944.
- [4] E. K. Fan, P. I. Kitov, W. F. Goebel, O. T. Avery, R. U. Lemieux, D. R. Bundle, D. A. Baker, V. Pozsgay, S. J. Danishefsky, J. R. Allen, E. J. Gordon, W. J. Sanders, L. L. Kiessling, L. E. Strong, S. D. Rosen, T. K. Kishimoto, M. A. Jutila, E. L. Berg, E. C. Butcher, P. J. Beck, R. Alon, C. R. Bertozzi, P. G. Wang, T. A. Fritz, W. H. Taylor, J. D. Esko, A. Heisner, P. Stehling, F. Oesch, T. J. Fralich, S. Itonori, Y. Ichikawa, K. J. Yarema, L. K. Mahal, R. Bruehl, E. C. Rodriguez, C. L. Jacobs, L. Strong, and E. Saxon, *Science* **2001**, 291, 2364-70.
- [5] A. Dove, *Nat. Biotechnol.* **2001**, 19, 913-7.
- [6] R. A. Dwek, *Biochem. Soc. Trans.* **1995**, 23, 1-25.
- [7] R. A. Dwek, *Chem. Rev.* **1996**, 96, 683-720.
- [8] K. A. Z. Siddiquee, P. T. Gunning, M. Glenn, W. P. Katt, S. Zhang, C. Schroeck, S. M. Sebti, R. Jove, A. D. Hamilton, and J. Turkson, *ACS Chem. Biol.* **2007**, 2, 787-98.
- [9] D. B. Werz, R. Ranzinger, S. Herget, A. Adibekian, C. W. von der Lieth, and P. H. Seeberger, *ACS Chem. Biol.* **2007**, 2, 685-91.

- [10] A. Adibekian, P. Stallforth, M-L. Hecht, D. B. Werz, P. Gagneux, and P. H. Seeberger, *Chem. Sci.* **2011**, 2, 337-44.
- [11] J. Cheng, *BMC Struct. Biol.* **2008**, 8, 18.
- [12] B. G. Davis, *Chem. Rev.* **2002**, 44, 579-601.
- [13] A. Varki, *Glycobiology* **1993**, 3, 97-130.
- [14] T. Angata and A. Varki, *Chem. Rev.* **2002**, 102, 439-69.
- [15] M. B. Jones, H. Teng, J. K. Rhee, N. Lahar, G. Baskaran, and K. J. Yarema, *Biotechnol. Bioeng.* **2004**, 85, 394-405.
- [16] W-D. Fessner, C. Goße, G. Jaeschke, and O. Eyrisch, *European J. Org. Chem.* **2000**, 125-32.
- [17] A. P. Corfield and M. Berry, *Trends Biochem. Sci.* **2015**, 40, 351-9.
- [18] H-J. Gabius, H-C. Siebert, S. André, J. Jiménez-Barbero, and H. Rüdiger, *ChemBioChem* **2004**, 5, 740-64.
- [19] H-J. Gabius, S. André, H. Kaltner, and H. C. Siebert, *Biochim. Biophys. Acta* **2002**, 1572, 165-77.
- [20] A. Bardosi, L. Bardosi, R. Lindenblatt, and H. Gabius, *Histochemistry* **1990**, 94, 285-91.
- [21] U. Schumacher, *Histochemistry* **1992**, 97, 95-99.
- [22] A. Ogura, A. Kurbangalieva, and K. Tanaka, *Glycoconj. J.* **2014**, 31, 273-9.
- [23] T-L. Hsu, S. R. Hanson, K. Kishikawa, S-K. Wang, M. Sawa, and C-H. Wong, *Proc. Natl. Acad. Sci. U. S. A.* **2007**, 104, 2614-9.
- [24] D. P. Praaning-van Dalen, A. M. de Leeuw, A. Brouwer, and D. L. Knook, *Hepatology* **1987**, 7, 672-9.
- [25] K.A.Deal, M.E.Cristel, and M. J. Welch, *Nucl. Med. Biol.* **1998**, 25, 379-85.

- [26] B. G. Davis and M. A. Robinson, *Curr. Opin. Drug Discov. Devel.* **2002**, 5, 279-88.
- [27] T. McCann and N. Kosaka, *Bioconjug. Chem.* **2010**, 21, 1925-32.
- [28] N. Yamazaki, S. Kojima, N. V. Bovin, S. André, S. Gabius, and H-J. Gabius, *Adv. Drug Deliv. Rev.* **2000**, 43, 225-44.
- [29] P. C. Calder, G. Dimitriadis, and P. Newsholme, *Curr. Opin. Clin. Nutr. Metab. Care* **2007**, 10, 531-40.
- [30] T. A. Springer, *Annu. Rev. Physiol.* **1995**, 57, 827-72.
- [31] J. C. McAuliffe and O. Hindsgaul in *Molecular and Cellular Glycobiology* (Eds.:M. Fukuda, O. Hindsgaul), Oxford University Press, Oxford, **2000**.
- [32] B. Ernst and J. L. Magnani, *Nat. Rev. Drug Discov.* **2009**, 8, 661-77.
- [33] V. Verez-Bencomo, V. Fernandez-Santana, E. Hardy, M. E. Toledo, M. C. Rodriguez, L. Heynngnezz, A. Rodriguez, A. Baly, L. Herrera, M. Izquierdo, A. Villar, Y. Valdes, K. Cosme, M. L. Deler, M. Montane, E. Garcia, A. Ramos, A. Aguilar, E. Medina, G. Torano, I. Sosa, I. Hernandez, R. Martinez, A. Muzachio, A. Carmenates, L. Costa, F. Cardoso, C. Campa, M. Diaz, and R. Roy, *Science* **2004**, 305, 522-5.
- [34] A. Phalipon, M. Tanguy, C. Grandjean, C. Guerreiro, F. Bélot, D. Cohen, P. J. Sansonetti, and L. A. Mulard, *J. Immunol.* **2009**, 182, 2241-7.
- [35] P. Simerska, H. Lu, and I.Toth, *Bioorg. Med. Chem. Lett.* **2009**, 19, 821-4.
- [36] A. F. G. Bongat, R. Saksena, R. Adamo, Y. Fujimoto, Z. Shiokawa, D. C. Peterson, K. Fukase, W. F. Vann, and P. Kovác, *Glycoconj. J.* **2010**, 27, 69-77.
- [37] C. Geraci, G. M. L. Consoli, E. Galante, E. Bousquet, M. Pappalardo, and A. Spadaro, *Bioconjug. Chem* **2008.**, 19, 751-8.

- [38] A. Miermont, H. Barnhill, E. Strable, X. Lu, K. A. Wall, Q. Wang, M. G. Finn, and X. Huang, *Chem Eur J* **2008**, 14, 4939-47.
- [39] O. Renaudet, L. BenMohamed, G. Dasgupta, I. Bettahi, and P. Dumy, *ChemMedChem* **2008**, 3, 737-41.
- [40] Q. Wang, S. A. Ekanayaka, J. Wu, J. Zhang, and Z. Guo, *Bioconjug. Chem.* **2008**, 19, 2060-7.
- [41] S. Bay, S. Fort, L. Birikaki, C. Ganneau, E. Samain, Y-M. Coïc, F. Bonhomme, E. Dériaud, C. Leclerc, and R. Lo-Man, *ChemMedChem* **2009**, 4, 582-7.
- [42] I. Bettahi, G. Dasgupta, O. Renaudet, A. A. Chentoufi, X. Zhang, D. Carpenter, S. Yoon, P. Dumy, and L. B. Mohamed, *Cancer Immunol. Immunother.* **2009**, 58, 187-200.
- [43] S. Ingale, M. A. Wolfert, T. Buskas, and G-J. Boons, *ChemBioChem* **2009**, 10, 455-63.
- [44] I. Jeon, K. Iyer, and S. J. Danishefsky, *J. Org. Chem.* **2009**, 74, 8452-55.
- [45] D. Lee and S. J. Danishefsky, *Tetrahedron Lett.* **2009**, 50, 2167-70.
- [46] P. Nagorny, H. K. Woo, Q. Wan, D. Lee, and S. J. Danishefsky, *J. Org. Chem.* **2009**, 74, 5157-62.
- [47] G. Ragupathi, P. Damani, G. Srivastava, O. Srivastava, S. J. Suchecke, Y. Ichikawa, and P. O. Livingston, *Cancer Immunol. Immunother.* **2009**, 58, 1397-1405.
- [48] J. Zhu, Q. Wan, D. Lee, G. Yang, M. K. Spassova, O. Ouerfelli, G. Ragupathi, P. Damani, P. O. Livingston, and S. J. Danishefsky, *J. Am. Chem. Soc.* **2009**, 131, 9298-9303.

- [49] C. L. Brooks, A. Schietinger, S. N. Borisova, P. Kufer, M. Okon, T. Hirama, C. R. Mackenzie, L-X. Wang, H. Schreiber, and S. V Evans, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, 107, 10056-61.
- [50] O. Renaudet, G. Dasgupta, I. Bettahi, A. Shi, A. B. Nesburn, P. Dumy, and L. BenMohamed, *PLoS One* **2010**, 5, e11216.
- [51] R. D. Astronomo and D. R. Burton, *Nat. Rev. Drug Discov.* **2010**, 9, 308-324.
- [52] F. Paillard, *Hum. Gene Ther.* **1999**, 10, 337-39.
- [53] P. Ehrlich, *Chemotherapy* **1960**, 1-615.
- [54] L. W. Seymour, *Adv. Drug Deliv. Rev.* **1994**, 14, 89-111.
- [55] M. Monsigny, P. Midoux, R. Mayer, and A. C. Roche, *Biosci. Rep.* **1999**, 19, 125-132.
- [56] G. Ashwell and J. Harford, *Annu. Rev. Biochem.* **1982**, 51, 531-54.
- [57] G. Y. Wu and C. H. Wu, *Adv. Drug Deliv. Rev.* **1998**, 29, 243-8.
- [58] A. David **2010**, *Isr. J. Chem.*, 50, 204-19.
- [59] Y. Liu, J. Liu, X. Pang, T. Liu, Z. Ning, G. Cheng, *Molecules* **2015**, 20, 2272-95.
- [60] P. Ma, S. Liu, Y. Huang, X. Chen, L. Zhang, and X. Jing, *Biomaterials* **2010**, 31, 2646-54.
- [61] G. Wilson, *J. Biol. Chem.* **1978**, 253, 2070-2.
- [62] R. J. Stockert, A. G. Morell, and H. Scheinberg, *Science* **1974**, 186, 365-366.
- [63] J. J. Lundquist and E. J. Toone, *Chem. Rev.* **2002**, 102, 555-78.

- [64] D. Benito-Alifonso, S. Tremel, B. Hou, H. Lockyear, J. Mantell, D. J. Fermin, P. Verkade, M. Berry, and M. C. Galan, *Angew. Chemie Int. Ed.* **2014**, 53, 810-4.
- [65] A. Ketkar-Atre, T. Struys, T. Dresselaers, M. Hodenius, I. Mannaerts, Y. Ni, I. Lambrichts, L. A. Van Grunsven, M. De Cuyper, and U. Himmelreich, *Biomaterials*, 35, 1015-24.
- [66] M. Mammen, S. K. Choi, and G. M. Whitesides, *Angew. Chemie Int. Ed.* **1998**, 37, 2754-94.
- [67] L. E. Kiessling, L. Getwicky, J. E. Strong, *Angew Chem Int Ed Engl* **2006**, 45, 2348-68.
- [68] C. Fasting, C. A. Schalley, M. Weber, O. Seitz, S. Hecht, B. Kokschi, J. Darnedde, C. Graf, E-W. Knapp, and R. Haag, *Angew. Chemie Int. Ed.* **2012**, 51, 10472-98.
- [69] P. I. Kitov and D. R. Bundle, *J. Am. Chem. Soc.* **2003**, 125, 16271-84.
- [70] A. Mulder, J. Huskens, and D. N. Reinhoudt, *Org. Biomol. Chem.* **2004**, 2, 3409-24.
- [71] C.A. Hunter and H.L. Anderson, *Angew.Chemie Int.Ed.***2009**, 48, 7488-99.
- [72] Y. C. Lee and R. T. Lee, *Acc. Chem. Res.* **1995**, 28, 321-7.
- [73] H. Staudinger and J. Fritsch, *Helv. Chim. Acta* **1922**, 5, 785-806.
- [74] R. J. Ellis, *Trends Biochem. Sci.* **2001**, 26, 597-604.
- [75] F. H. Silver, in *Mechanical Properties of Biological Macromolecules*, Springer, US, **2006**, 168-180.
- [76] H-X. Zhou, G. Rivas, and A. P. Minton, *Annu Rev Biophys* **2008**, 37, 375-97.
- [77] B. Seal, *Mater. Sci. Eng. Reports* **2001**, 34, 147-230.

- [78] E. Mattia and S. Otto, *Nat. Nanotechnol.* **2015**, 10, 111-9.
- [79] D. Ma, X. Hu, H. Zhou, J. Zhang, and Y. Qian, *J. Cryst. Growth* **2007**, 304, 163-8.
- [80] A. V. Pinheiro, D. Han, W. M. Shih, and H. Yan, *Nat. Nanotechnol.* **2011**, 6, 763-72.
- [81] M.P. Kosloski, R. D. Miclea, and S. V Balu-Iyer, *AAPS J* **2009**, 11, 424-31.
- [82] D. W. Park, S. Haam, T. G. Lee, H.-S. Kim, and W.-S. Kim, *J. Biomed. Mater. Res.* **2004**, 71A, 497-507.
- [83] A. M. Sinclair and S. Elliott, *J Pharm Sci* **2005**, 94, 1626-35.

Carbohydrate synthesis

2.1 Introduction

2.1.1 Structure and nomenclature of carbohydrates

As their name suggests, carbohydrates are hydrated forms of carbon. Unsurprisingly their empirical formula is therefore $C_m(H_2O)_n$ (with some exceptions, such as deoxyribose $C_5H_{10}O_4$). Sugar structures are characterised by a series of hydroxyl groups and one aldehyde (aldoses) or ketone (ketoses). These structure exists in an equilibrium between open-chain and closed ring forms.^[1] The ring forms from the reaction of the aldehyde (or ketone) with a hydroxyl group, generating a hemiacetal (or hemiketal) (Fig. 2.1).

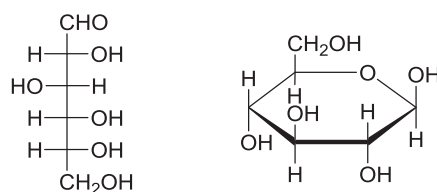


Figure 2.1 Open-chain (Fisher projection) and closed ring (Haworth projection) representations of D-Glc.

For hexoses, the chain closes in a 5 or 6 membered ring, called furanose or pyranose rings respectively. Depending on the 3D geometry of the molecule, the furanose ring can be found in 2 conformations, twist and envelope, while the

pyranose can be found in 5 different conformers: chair, boat, skew, half-chair and envelope. The pyranose ring chair can also have ${}^1C_4$ or ${}^4C_1$ conformation (Fig. 2.2).^[1]

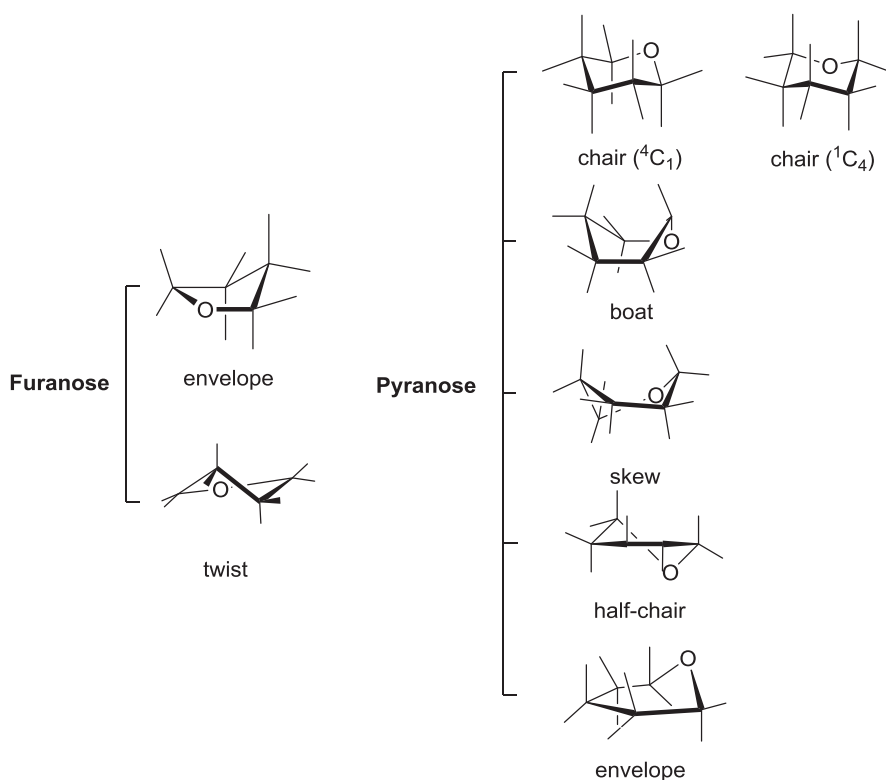


Figure 2.2 Furanose and pyranose conformers.

All the different conformers are in equilibrium with each other and with the open-chain form. These forms can interconvert into one another if the appropriate energy requirements are met. The abundance of one conformer in solution depends mainly on the steric and stereoelectronic effects dictated by the substituents of the hydroxyl groups (1,3 diaxial interactions, dipoles, hydrogen bonding and the anomeric effect), but also on the nature of the solvent, temperature and pH.

Adding complexity to the structural analysis of carbohydrates is their chirality: sugars are divided into D- (dextrorotatory) and L- (laevorotatory) forms depending on their ability to rotate plane-polarised light clockwise or counter-clockwise, which is the result of the sum of all their chiral centres. The configuration of the chiral centres determines the identity of the sugars. Stereoisomers at any one chiral centre are called epimers (e.g. glucose and galactose, Fig. 2.3) while stereoisomers at the anomeric centre are called anomers. Every saccharide possesses two different anomers, α and β (Figure 2.3).

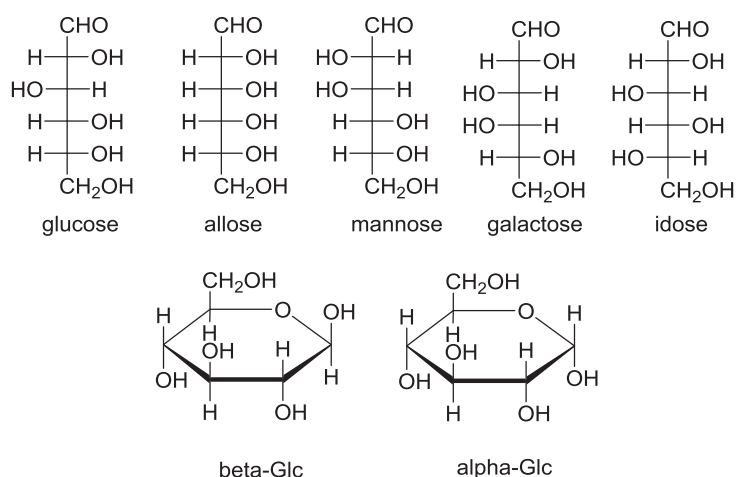


Figure 2.3 Epimers (top) and anomers (bottom) of Glc.

Drawing the structure of oligosaccharides (OS) becomes more difficult and time consuming with the increase in sugar number, types and connections between them. Symbolic diagrams are widely used to make the interpretation of complex structures clearer to other scientists. Unfortunately, there is no general agreement between researchers, leading to inconsistency and confusion. There are multiple challenges to face: we first need to take into account the constituent monosaccharides alone to find a way to represent all of their features. Symbols

that would be suitable to represent monosaccharide building blocks should display more information than simply the sugar identity, such as ring size, configuration, anomericity and substitution. These symbols should then be connected in a logical and unambiguous way to represent the linkage anomericity and position. A standard system for drawing complex OS has not yet been established, although two main schools of thought have emerged, after Kornfield's first proposal.^[2] In 2009 Harvey et al.^[3] devised a strategy based on using black and white symbols of different shapes (square, circle, star, diamond) with very basic patterns (dot, cross, filling, no filling) (Table. 2.1) and connected them with dashed or solid lines, representing β or α linkages. The linkage position was denoted by the angle of the line linking the sugar units.

Sugar	Basic	N-Ac	NH ₂	(NH ₂) ₂	Deoxy	Acid
Hexose						
Glucose	□ □ Glucose	■ ■ GlcNAc	▣ ▣	▤ ▤	▥ ▥	▧ ▧
Galactose	◇ ◇ Galactose	◆ ◆ GalNAc	◊ ◊	◈	◉ ◉ Fucose	◊ ◊
Mannose	○ ○ Mannose	● ●	◐ ◐	◑ ◑	◒ ◒	◓ ◓
Idose	⊕	⊕	⊕	⊕	⊕	⊕
Unknown or unspecified						
Pentose	⬡	⬢	⬣	⬤	⬥	-
Hexose	⬦	⬧	⬨	⬩	⬪	-
Pentose						
Xylose	△ △ Xylose	▲ ▲	▴ ▴	-	▵ ▵	-
Ribose	▽	▼	▾	-	▿	-
Arabinose	▷	▶	▸	-	▹	-
Sialic acids						
Neu5Ac	-	★ ★ Neu5Ac	-	-	-	-
Neu5Gc	-	☆ ☆ Neu5Gc	-	-	-	-
Miscellaneous						
Inositol	☆	-	-	-	-	-
Unsat-GlcA	⊠	-	-	-	-	-
Tyvelose	⊙	-	-	-	-	-

Table 2.1 Harvey symbolic nomenclature of OS.

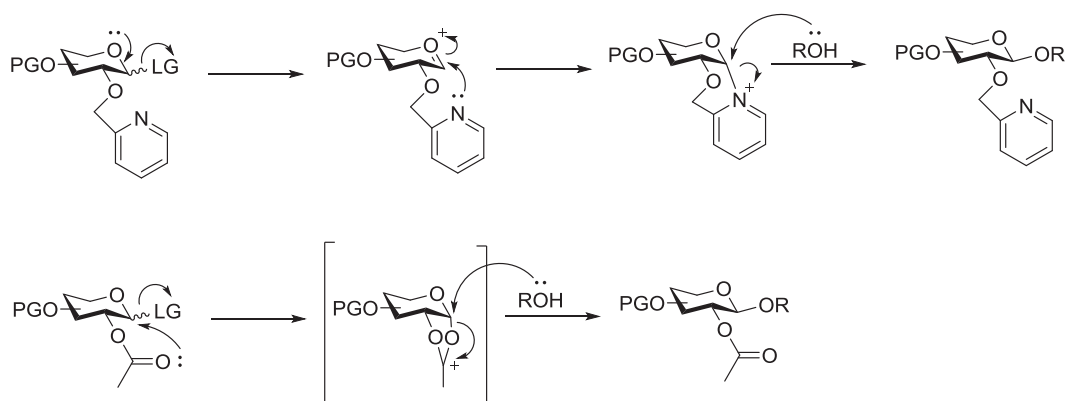
This proposal provoked the immediate reaction of the authors of 'Essentials of Glycobiology',^[4] who suggested instead the adoption of coloured symbols and moved the focus more towards the glycan composition and biosynthesis rather than providing information about the glycan structure and linkages. In a subsequent reply,^[5] Harvey and co-workers underlined the advantages of their approach, whose main strong points are the simplicity of drawing the symbols by hand and interpreting structures in an unambiguous way with black and white printed schemes. Not to mention the added value of representing the actual structures respecting their anomericity and connectivity. Intrigued by this still open debate, we decided to adopt the approach proposed by Harvey and colleagues, since having a clear representation of the OS and their connectivity will help our description of their structure and synthesis.

2.1.2 Synthesis of carbohydrates

In order to gain a deeper understanding of glycan diversity and functions it is essential to have access to structurally defined OS in adequate quantity and high purity. Approaches to preparing complex sugars in limited time and large amounts are lagging behind. Progress has been hampered by the need to isolate some starting materials from natural sources, lengthy syntheses or un-scalable enzymatic approaches.^[6] At first sight, glycosylation reactions do not look more complicated than other reactions. Indeed, a glycosyl donor is reacted with an acceptor in the presence of a promoter to yield the desired compound. Unfortunately, a glycosylation reaction is more likely to give a rainbow of products that would not surprise a carbohydrate chemist. Each glycosylation is a unique reaction that requires a lot of effort to optimise conditions such as equivalents,

solvent, temperature, time, and purification, in order to obtain an acceptable outcome. Although the formation of by-products is often mentioned, there are almost no publications reporting failed synthesis or describing the undesired sugar products in detail. This information would indeed be very useful but is often unavailable, as during the unwinding of the multistep synthesis of a saccharide, the efforts applied to obtain the product desired do not leave time and resources to investigate the formation of by-products. Nonetheless, common side reactions of glycosyl donors are known^[7] and can be listed as hydrolysis, elimination, intermolecular aglycon transfer and rearrangements.

A pivotal role is played in OS synthesis by protecting groups (PGs). The selective formation of a glycosidic bond, of desired anomericity,^[8] relies heavily on the appropriate choice of PGs for both donor and acceptor. Other than preventing hydroxyl groups which should not be involved in the glycosylation from reacting, PGs modulate the reactivity of the sugars through their stereo and electronic effects. For example, participating PGs tune the formation of the new bond by stabilising the oxonium intermediate or assisting the departure of the leaving group (LG) (Scheme 2.1).



Scheme 2.1 Mechanism of action of participating PGs.

Conformation-constraining PGs restrict the flexibility of the sugar ring to a certain conformation, dictating the formation pathway of the new glycosidic bond.

New synthetic strategies to produce carbohydrates more efficiently are reactivity based one-pot convergent approaches. Pioneering this field was the Fraser-Reid armed-disarmed methodology,^[9] developed further by the groups of Ley,^[10] Wong^[11] and Huang.^[12] Polymer supported OS synthesis^[13-15] was developed following the footsteps of the success of solid-phase peptide synthesis.^[16] One less well-known solid-phase based approach is the surface-tethered iterative carbohydrate synthesis (STICS), based on highly functionalised porous gold surfaces used as support.^[17] Moving to homogeneous phase reactions, fluorinated soluble support strategies have been developed.^[18-21] A fluorinated hydrazide linker-tag was reacted with ketone-functionalised OS, and the crude mixture filtered through a fluorous solid-phase extraction column to selectively capture the fluorous-tagged material. The final product would later be recovered by elution with the appropriate solvent. To conclude with one final example, a new strategy based on ionic liquids as soluble functional support has also been investigated.^[22-25]

2.1.3 Characterisation of carbohydrates

A combination of various analytical techniques are used simultaneously to obtain full characterisation of carbohydrates. While mass spectroscopy and elemental analysis give information on the elemental composition and weight of samples, they provide limited clues to their structures.^[26] Infra-red spectroscopy (IR) probes the presence of characteristic formations (functional groups, bonds etc.) in the

molecules but does not suggest how these are connected. Nuclear magnetic resonance (NMR) is the characterisation technique of choice for the structural analysis of carbohydrates. Coupling constants between protons, 2D proton-proton and proton-carbon correlations unravel the relationships between atoms in the structure, the anomeric configuration, their conjugation and even spatial arrangement. To assign the anomeric configuration we cannot always rely only on the value of the coupling constant of the anomeric proton; in some cases, such as for Man or Neu derivatives, we need to look at the coupling constants between the anomeric carbon $^{13}\text{C}(1)$ and protons $\text{H}(1)$ $^1\text{J}[^{13}\text{C}-\text{H}(1)]$ (Fig. 2.4). In pyranoses these are about 160 Hz in the β -anomers (axial $\text{H}1$) and 170 Hz in the α -anomers (equatorial $\text{H}1$).^[27] The electronegativity of the substituent at $\text{C}1$ (but not at $\text{C}2$) affects the value of the coupling constant $^1\text{J}[^{13}\text{C}-\text{H}(1)]$. These two values are different because the equatorial $\text{H}1$ is closer to the lone pair of the oxygen in the ring than the axial $\text{H}1$, resulting in a larger coupling constant.^[28-33]

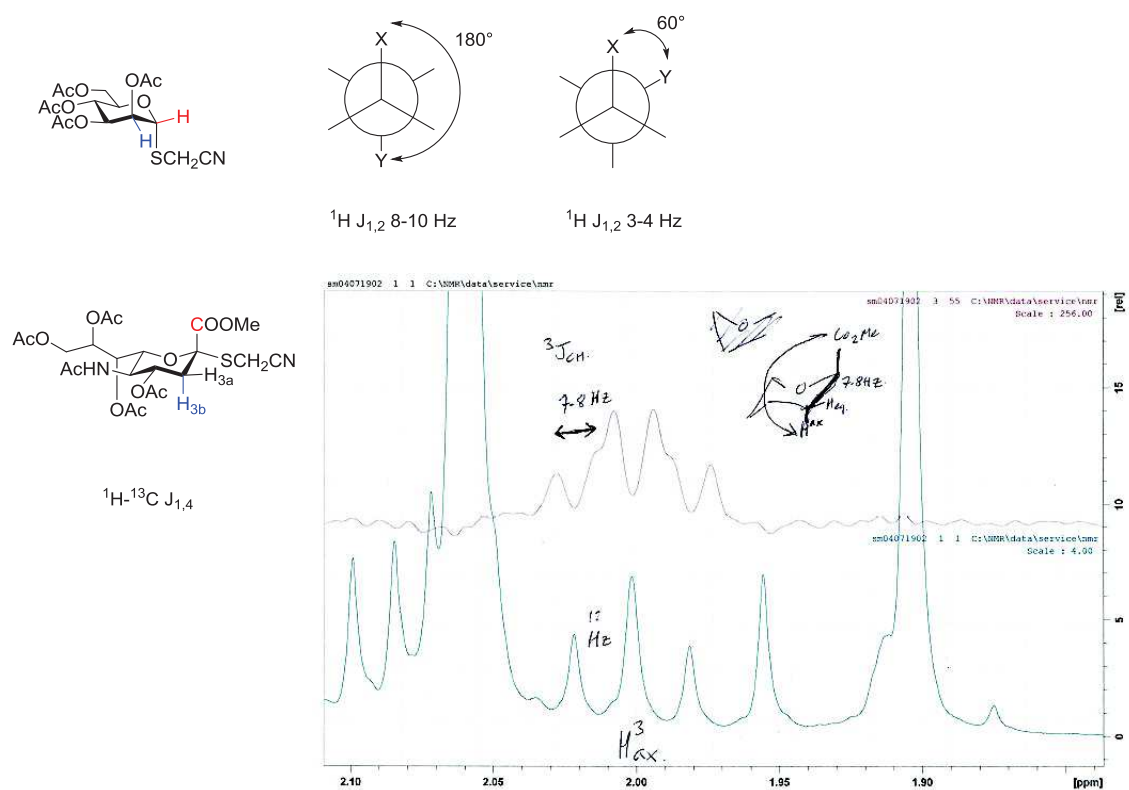
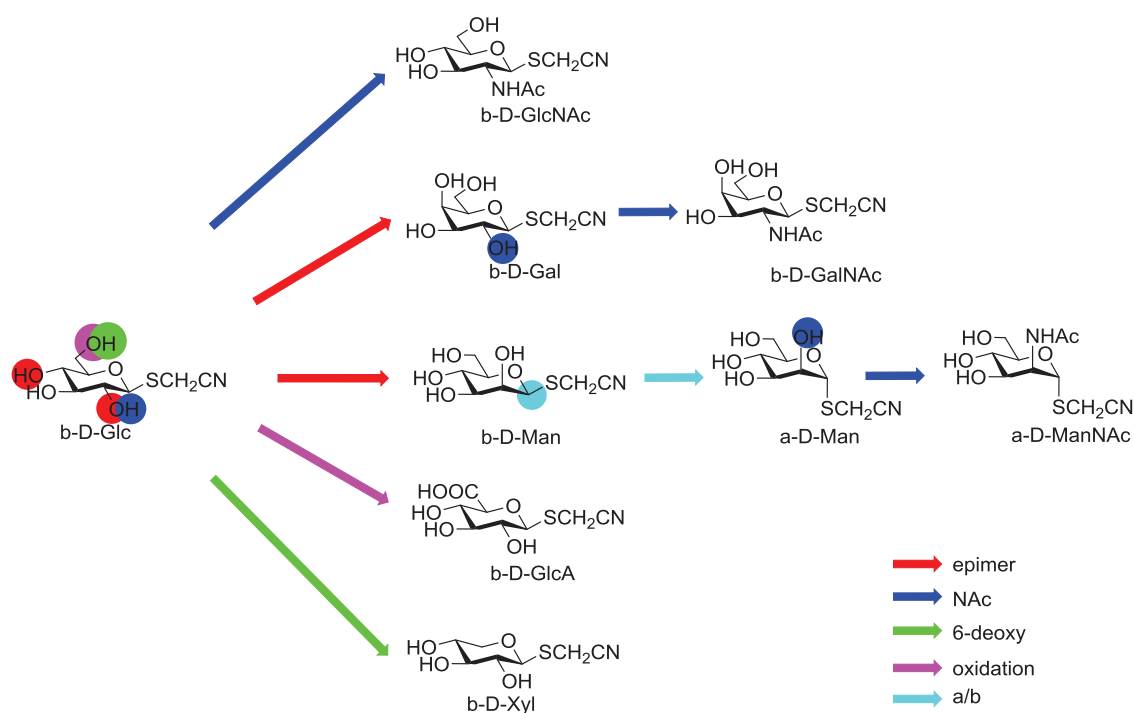


Figure 2.4 Assignment of the anomeric configuration of carbohydrates by NMR.

NMR studies, combined with molecular dynamic simulations, can be used to determine the conformation of OS. This would not be achievable by X-ray crystallography alone because of the sugars' dynamic fluctuations.

2.2 Glycomap sugar library

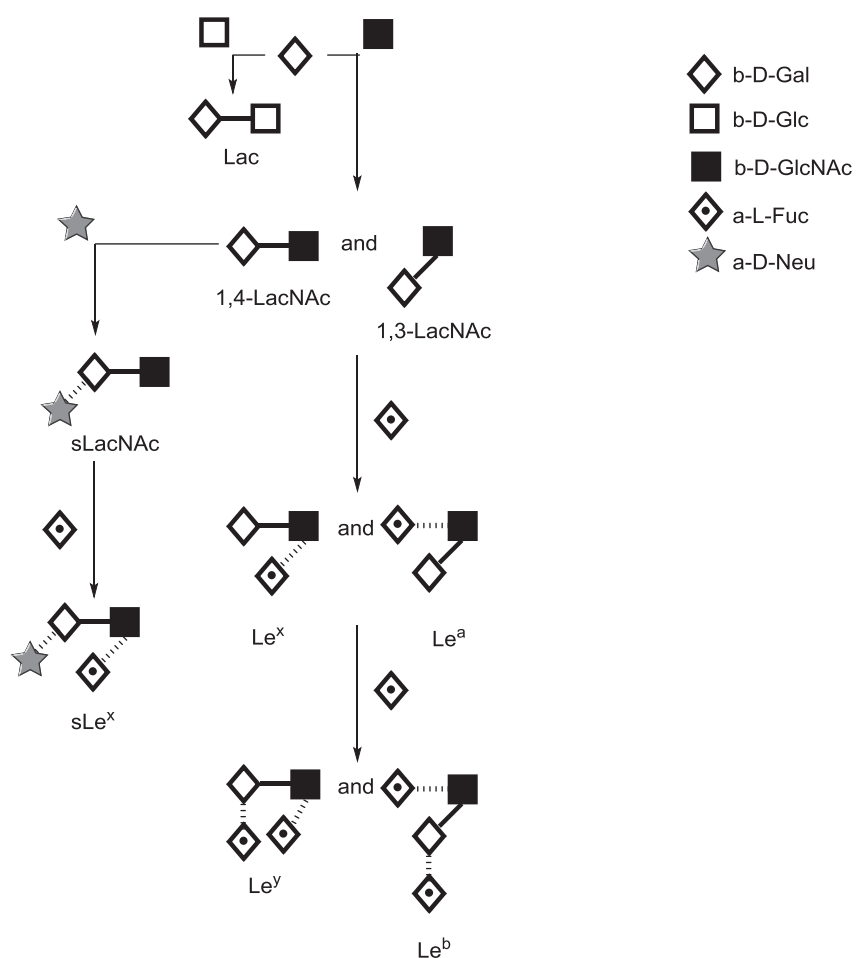
The sugars composing our library have been chosen following a dual approach: first, through the rational and systematic chemical modification of β -D-Glc (Scheme 2.2); and secondly taking into account their biological relevance. This means that the derivatisations introduced have been selected from some of the possible chemical modifications of carbohydrates in order to yield sugars of biological significance.



Scheme 2.2 Glycomap monosaccharides: from β -D-Glc a logical sequence of single chemical modifications leads to the exploration of several features of carbohydrates such as different anomers, epimers, N-Acetylated and oxy/deoxy derivatives.

Glucose (Glc), N-acetyl glucosamine (GlcNAc), galactose (Gal), N-acetyl galactosamine (GalNAc), mannose (Man), xylose (Xyl), fucose (Fuc), glucuronic acid (GlcA), iduronic acid (IdoA) and sialic acid (Neu) are the ten monosaccharides constituting the building blocks of the entire mammalian glycome.^[34,35] These ten sugars found in humans can be assembled into more than 15 million possible tetrasaccharides. Five out of ten of the mammalian monosaccharides (Gal, Man, GlcNAc, GlcA and Xyl) can be obtained from Glc as common precursor, by introducing one single modification. The nature of these modifications is diverse and represents only few of all the possible chemical manipulations that can be performed on carbohydrates. From this first set of monosaccharides a new generation of sugars will emerge by introduction of a

second modification, which brings into the picture one more basic mammalian monosaccharide (GalNAc). The real potential of monosaccharides can be truly deconvoluted only once we explore their multipotent ability to combine in more complex structures. From two of the monosaccharides previously described, we will explore the route leading towards the benchmark compound for selectin recognition, sLe^x[36] (Scheme 2.3).



Scheme 2.3 Glycomap oligosaccharides: from the assembly of two monomers we built more complex sugars. The progressive decoration of the core disaccharide 1-3 or 1-4,LacNAc leads to a series of fucosylated and sialylated derivatives of biological relevance.

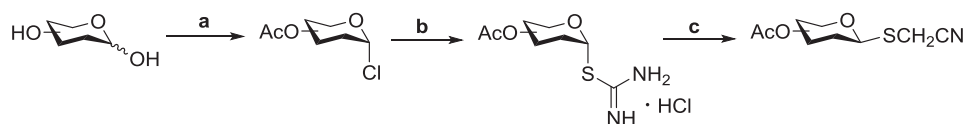
Lactose (Lac), is the first intermediate encountered on our way to sLe^x, formed by the combination of Gal and Glc in a 1,4- β linkage. If the acceptor is the N-acetylated derivative of Glc (GlcNAc) we end up with (1,4)-LacNAc, which constitutes the core of our most complex tetrasaccharide. Taking it one step further, the combination of (1,4)-LacNAc, Fuc and Neu yields a series of biologically interesting saccharides, amongst which we can finally place sLe^x. The biological relevance of these carbohydrates has already been introduced in Chapter 1 and will be described more in depth in Chapter 6. S-cyanomethyl group (SCM) has been installed in the anomeric position of this library of carbohydrates as it has good stability during chemical and enzymatic transformations.^[37] The SCM group can be easily interconverted into the reactive 2-imino-2-methoxyethyl (IME) form at a later stage.

2.3 Carbohydrate synthesis

2.3.1 Monosaccharides

The synthesis of the monosaccharides has been performed following a basic common strategy that has been tailored for each of them. In this section we will describe in detail the differences in their synthetic protocols.

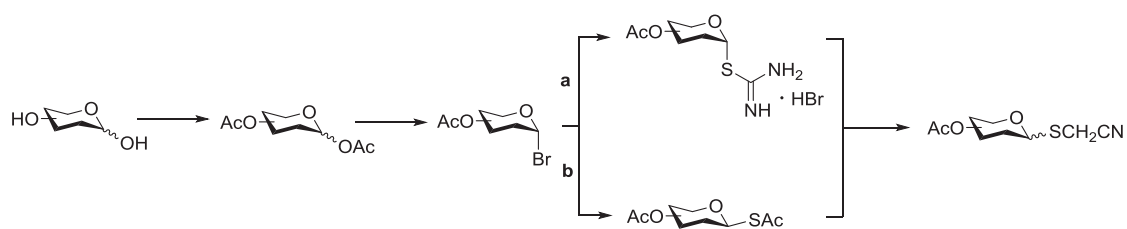
GlcNAc **2.3** and GalNAc **2.45** have been prepared following a 3-step synthesis (Scheme 2.4).



Scheme 2.4 Reagents and conditions: a) AcCl, r.t.; b) thiourea, acetone, 80 °C; c) Na₂S₂O₅, K₂CO₃, chloroacetonitrile, H₂O:acetone (1:1), r.t..

The fully deprotected starting material was peracetylated and simultaneously activated as 1- α -Cl derivative over 2 days in pure AcCl. While GlcNAc was obtained in 88% yield after re-crystallisation from cold Et₂O, GalNAc did not benefit from the re-crystallisation and was isolated in only 71% yield. The chloro-derivative was transformed into the thiouronium salt by refluxing the sugar in acetone with thiourea for 1 or 2 h. The precipitate formed was washed with cold EtOH giving GalNAc in 63% yield while GlcNAc-thiouronium was obtained in near quantitative yield. In the final step the thiouronium salt was reacted with chloroacetonitrile overnight, and the desired product was isolated in 54% yield for GalNAc (**2.45**) and 87% yield for GlcNAc (**2.3**). As a final remark we would like to emphasise that these 3-step syntheses yielded the desired products in good yields, without any chromatographic purification, in scalable conditions.

Most of the other monosaccharides (Gal **2.42**, ManNAc **2.59**, Glc **2.55**, α - and β -Man, **2.51** and **2.53**, Xyl **2.62** and Fuc **2.63**) we synthesised following a very similar strategy to the one described above but going through a bromo- instead of a chloro- intermediate as the latter was too unstable to be isolated (Scheme 2.5).

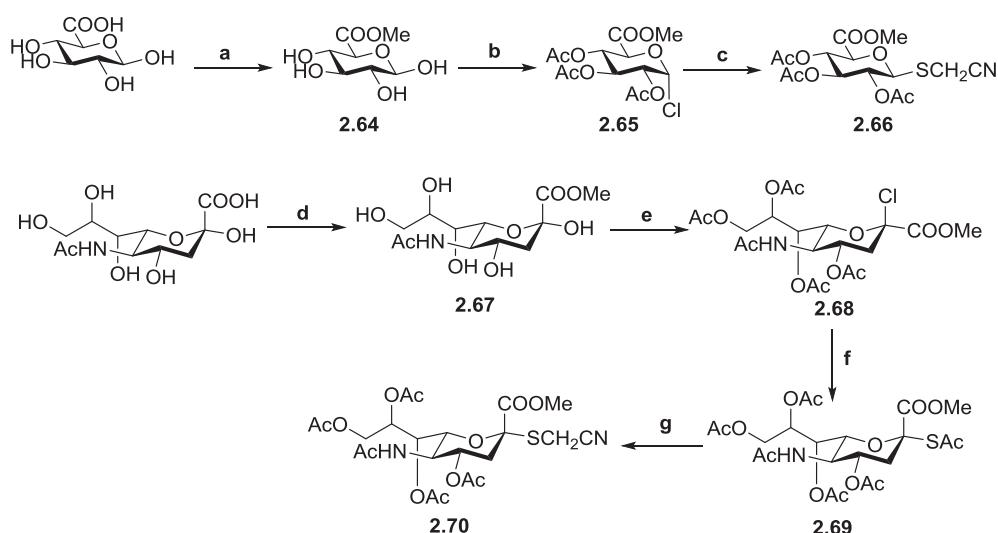


Scheme 2.5 Synthesis of monosaccharides through bromo-intermediates. The diversification of the penultimate intermediate defines the anomericity of the final product.

Deprotected starting materials were first peracetylated with Ac_2O and DMAP for 1 or 2 h before being recovered with EtOAc and purified by an aqueous work-up. ManNAc showed resistance to being transferred into the organic phase, and a very polar IPA:DCM (1:4) solvent mixture was required in order to effect the extraction. Peracetylation of Fuc followed a very different route: the sugar was first suspended in DCM before DMAP, triethylamine and Ac_2O were added at 0°C , and the reaction stirred overnight at r.t.. Regardless their different protocols, all peracetylated sugars were obtained in quantitative yields. 1- α -Br has been introduced in all sugars by treating the peracetylated materials with HBr and Ac_2O for about 1 h at r.t.; only for Glc this reaction occurred simultaneously with the peracetylation. Although ManNAc did not require any aqueous workup, its derivative was obtained in almost quantitative yield as for most of the other sugars, with the exception of Man and Gal which have been isolated in 52% and 68% yield respectively. At this point α - and β -Man had to be differentiated: α -Man, together with most of the other sugars, proceeded to the thiouronium hydrobromide intermediates (Scheme 2.5, route a); β -Man instead formed the S-acetyl intermediate (Scheme 2.5, route b). The thiouronium derivatives were obtained by refluxing the bromoglycosides in acetone (or ACN) with thiourea for

1 or 2 h. Only ManNAc (60% yield) and Gal (quantitative) were isolated, the others converted directly into S-cyanomethyl compounds. Thiouronium bromides have been reacted with chloroacetonitrile for 12 h (or just 1 h for ManNAc) at r.t. and isolated after work-up (69% **2.59**, 65% **2.55**, 63% **2.62**, 79% **2.42**, 27% **2.51**). While for ManNAc and Xyl the aqueous work-up was sufficient, to obtain pure products, Glc was re-crystallised from hot Et₂O, α -Man was purified by column chromatography, and Gal was precipitated with water and washed with cold Et₂O. Finally Fuc, which reacted with chloroacetonitrile in the presence of triethylamine in just 1 h, was quickly filtered over a silica plug to give the pure product **2.63** in 94% yield. The 1-SAc derivative of β -Man was prepared from the bromide intermediate by overnight reaction with KSAc, formed in situ from thioacetic acid and potassium tert-butoxide. By comparison, reaction with pre-formed KSAc resulted in very poor yield, even after multiple additions of KSAc. The product was finally obtained in 44% yield but required purification by flash column chromatography. Conversion of the S-acetyl β -Man into the S-cyanomethyl product was achieved after deprotection of the thiol at -40 °C with NaOMe, and reaction with chloroacetonitrile and DIPEA for 4 h. After another chromatographic purification the product **2.53** was isolated in 50% yield.

The last two monosaccharides prepared, GlcA **2.66** and Neu **2.70**, required the derivatisation of their free carboxyl group into the methyl ester before the strategies mentioned above could be applied successfully (Scheme 2.6).



Scheme 2.6 Reagents and conditions: a) DowexH⁺, MeOH, r.t., 94%; b) AcCl, AcOH, r.t., 77%; c) i. thiourea, acetone, 80 °C; ii. Na₂S₂O₅, K₂CO₃, chloroacetonitrile, H₂O:acetone (1:1), r.t., 23%; d) DowexH⁺, MeOH, r.t., 86%; e) AcCl, AcOH, r.t., quant.; f) KSac, DCM, r.t., 59%; g) i. NaOMe (25% in MeOH); ii. DIPEA, chloroacetonitrile, DCM, r.t., 72%.

The free carboxylic acid functionality was stirred in MeOH for 3 h with DowexH⁺ before the methyl ester was isolated in 94% yield for GlcA and 86% yield for Neu after a rapid filtration on silica plug. Peracetylation and 1-Cl derivatisation were performed simultaneously by reaction with AcCl in AcOH. The compound **2.68** was obtained in quantitative yield after 1 h, while GlcA could be isolated only in 77% after reacting for 22 h. The target GlcA-acetylated derivative compound **2.66** was obtained from 1-O-Me-per-O-Ac-GlcAAc **2.65** in a 2-step one-pot fashion (23% yield), by first transforming the chloro-derivative into the thiouronium salt refluxing with thiourea, and then stirring in chloroacetonitrile overnight at r.t.. NeuAc derivative compound **2.70** arrived at its final stage through the S-acetyl derivative, formed by reaction with KSac overnight at r.t. followed by purification by flash column chromatography (59% yield). The thiol was deprotected with NaOMe at -40 °C and reacted with chloroacetonitrile overnight to give the final product **2.70** in 72% yield after chromatographic purification.

2.3.2 Disaccharides

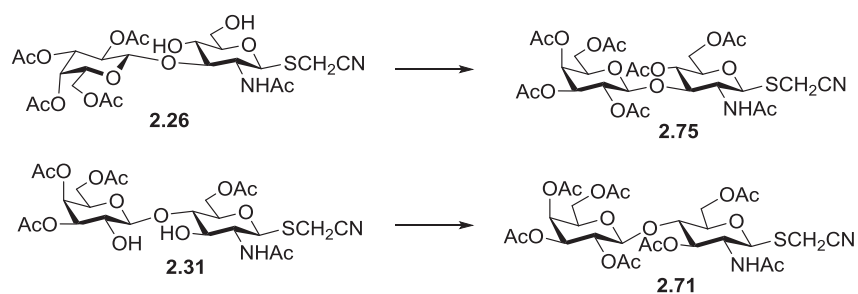
Synthesis of Lac **2.48** was performed using lactose as the starting material, via an analogous strategy to that used for the monosaccharides (Scheme 2.7).



Scheme 2.7 Reagents and conditions: a) Ac₂O, NaOAc, 140 °C, quant.; b) HBr (33% in AcOH), Ac₂O, r.t., 91%; c) i. thiourea, acetone, 80 °C; ii. Na₂S₂O₅, K₂CO₃, chloroacetonitrile, acetone, r.t., 63%.

Lactose was peracetylated in the presence of Ac₂O assisted by NaOAc at 140 °C for 5 h, giving the lactose-octaacetate in quantitative yield. Bromination of the anomeric position with HBr and Ac₂O led to the α-bromide product in 91% yield. The final step towards the S-cyanomethyl derivative was performed through the thiouronium salt (formed by refluxing with thiourea in acetone), then reacting with chloroacetonitrile in. After chromatographic purification the product **2.48** was obtained in 63% yield.

Preparation of (1,3)- and (1,4)-LacNAc (**2.75** and **2.71**) was performed in one step by peracetylation of the compounds **2.26** and **2.31** (Scheme 2.8), the synthesis of which is described later in this chapter (sections 2.3.3 and 2.3.4).



Scheme 2.8 Synthesis of (1,3)- and (1,4)-LacNAc. The two disaccharidic cores of the Lewis oligosaccharides are isolated as peracetylated analogues.

For the peracetylation of **2.26** the best conditions proved to be with AcCl and pyridine, which gave the product **2.75** in 89% yield after just 1 h reaction time. For the synthesis of **2.71** however, it was best was to use Ac₂O and DMAP, even though the final yield was just 32%. Both transformations did not require chromatographic purification.

2.3.3 Trisaccharides

The convergent synthesis of the Lewis oligosaccharides (Fig. 2.5) was previously designed in our group.^[38] The choice of protecting groups, donors and acceptors, and the assembly of the glycosidic building blocks have been thoroughly described. We synthesised Le^x **2.24** and Le^a **2.29** (together with Le^b **2.39** and Le^y **2.33**, described in section 2.3.4 of this chapter) reproducing the same protocols but including some minor improvements.

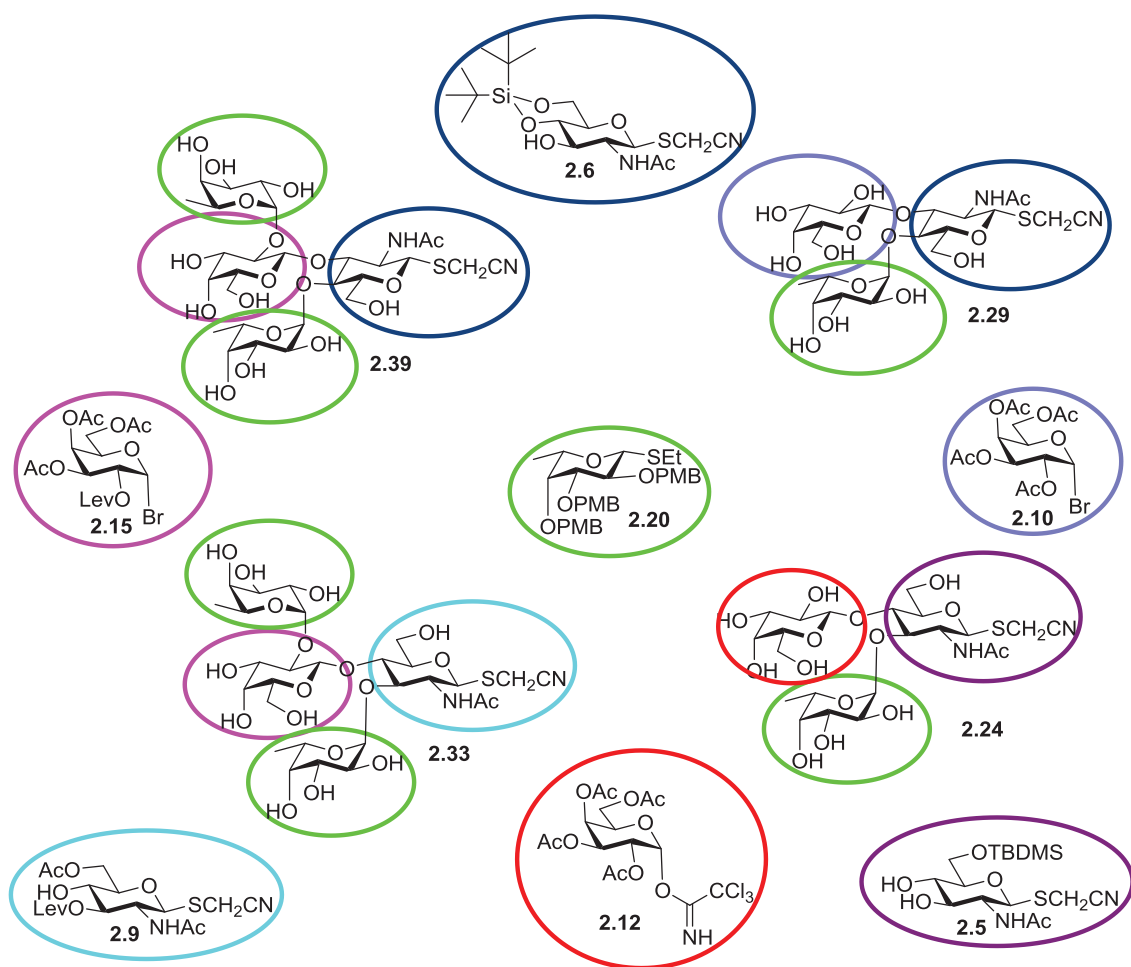
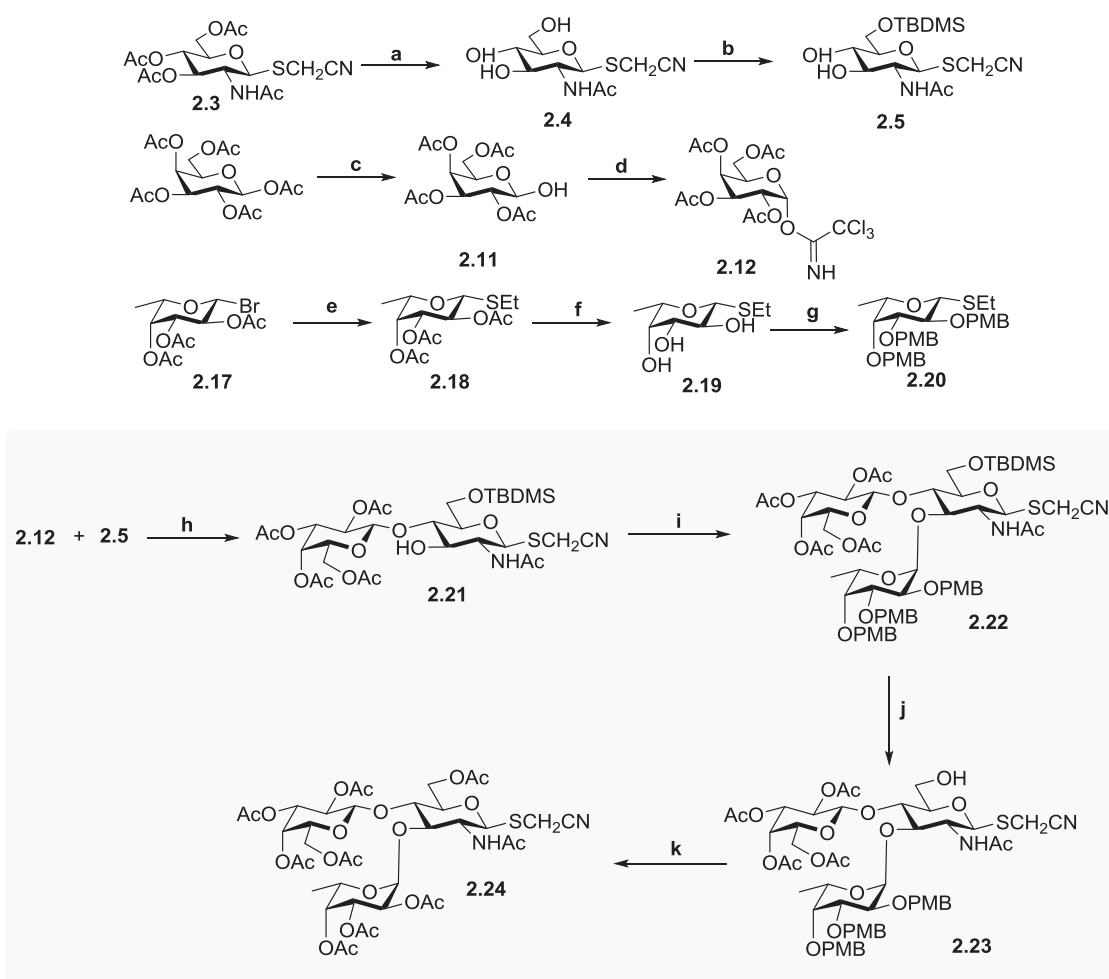


Figure 2.5 Convergent synthesis of Lewis OS. The four OS part of the Lewis family are composed by the respective monosaccharidic building blocks. Donors and acceptors for each OS are coded with same colour in the final products and as reagents.

Le^x **2.24** was synthesised in 9 steps from three precursors (**2.5**, **2.12**, **2.20**) through two glycosylation reactions (Scheme 2.9).

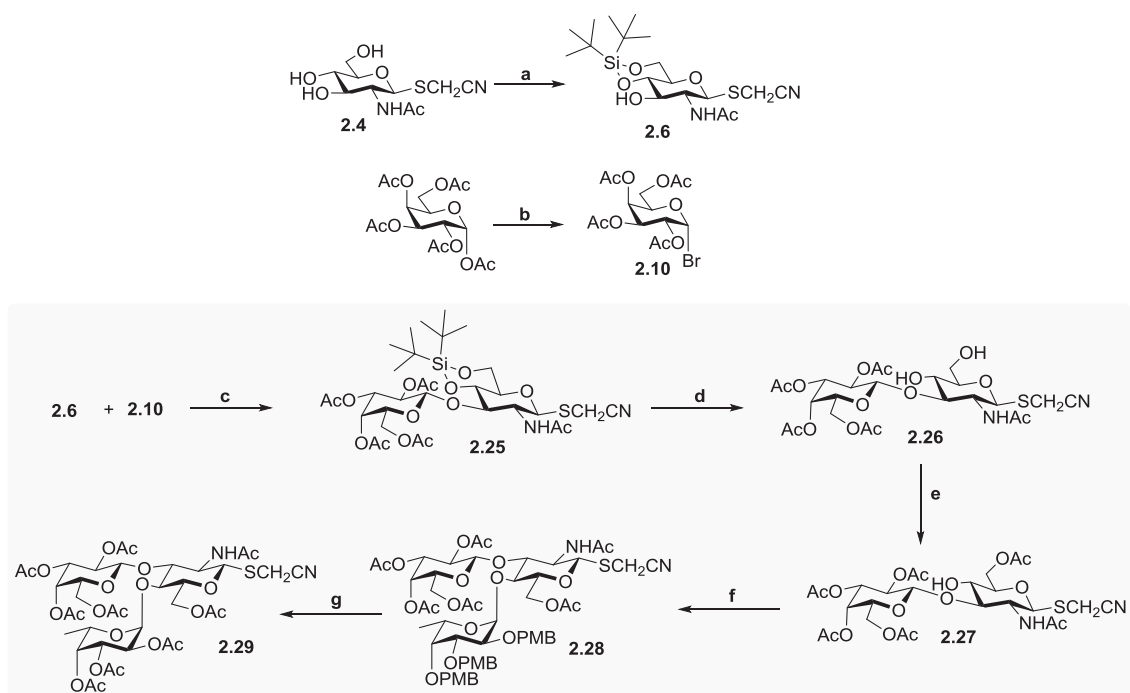


Scheme 2.9 Reagents and conditions: a) NEt_3 , MeOH, r.t., quant.; b) imidazole, TBDMS-Cl, DMF, r.t., 67%; c) hydrazine acetate, DMF, 50 °C, quant.; d) trichloroacetonitrile, DBU, DCM, r.t., 73%; e) i. thiourea, ACN, 85 °C; ii. triethylamine, ethylbromide, ACN, r.t., 75%; f) NaOMe (25% in MeOH), MeOH, r.t., quant.; g) TBAI, NaH, PMB-Cl, DMF, r.t., 49%; h) Et_2OBF_3 , DCM, MS, -41 °C, 42%; i) **2.20**, CuBr_2 , TBAB, DCM:DMF (1:1), MS, r.t., 79%; j) $\text{NEt}_3 \cdot 3\text{HF}$, THF, r.t., 84%; k) i. CAN; ii. Ac_2O , pyridine, DMAP, ACN:H₂O (9:1), r.t., 59%.

Precursor **2.5** was prepared from **2.3** by deprotection with NEt_3 , followed by reaction of the 6-OH with TBDMS-Cl and imidazole. Compound **2.12** was obtained from the activation of the deprotected anomeric position of pentaacetyl galactose with trichloro acetonitrile and DBU. Precursor **2.20** was synthesised from the compound **2.17** in 3 steps. First, the bromide was converted into the ethylthio derivative by reacting with ethylbromide and triethylamine, via the

formation of the thiouronium salt (refluxing with thiourea in ACN). The acetyl groups were removed with NaOMe in MeOH and replaced with paramethoxybenzyl (PMB) groups by reacting the free hydroxyl groups with tetrabutylammonium iodide (TBAI), NaH and PMB-Cl. Finally, with all our precursors in hand, the two glycosylation steps took place one after the other. The LacNAc core was first formed by reacting the donor **2.12** with the acceptor **2.5**. After that the disaccharide was fucosylated with the donor **2.20**. The first glycosylation was performed at low temperature with Et₂OBF₃ and quenched with triethylamine after 1h. Dry conditions were adopted, flame-drying the glassware used and adding activated molecular sieves. Unfortunately after the chromatographic purification the product **2.21** was obtained in only 42% yield. The second glycosylation, which utilised CuBr₂ and TBAB as the activation agent, was performed overnight in the dark. After flash chromatography the product **2.22** was isolated in 79% yield. The TBDMS group was removed from trisaccharide **2.22** using NEt₃·3HF, and the PMB groups with CAN, after which the trisaccharide was peracetylated using Ac₂O, pyridine and DMAP to obtain the final compound **2.24** in 59% yield.

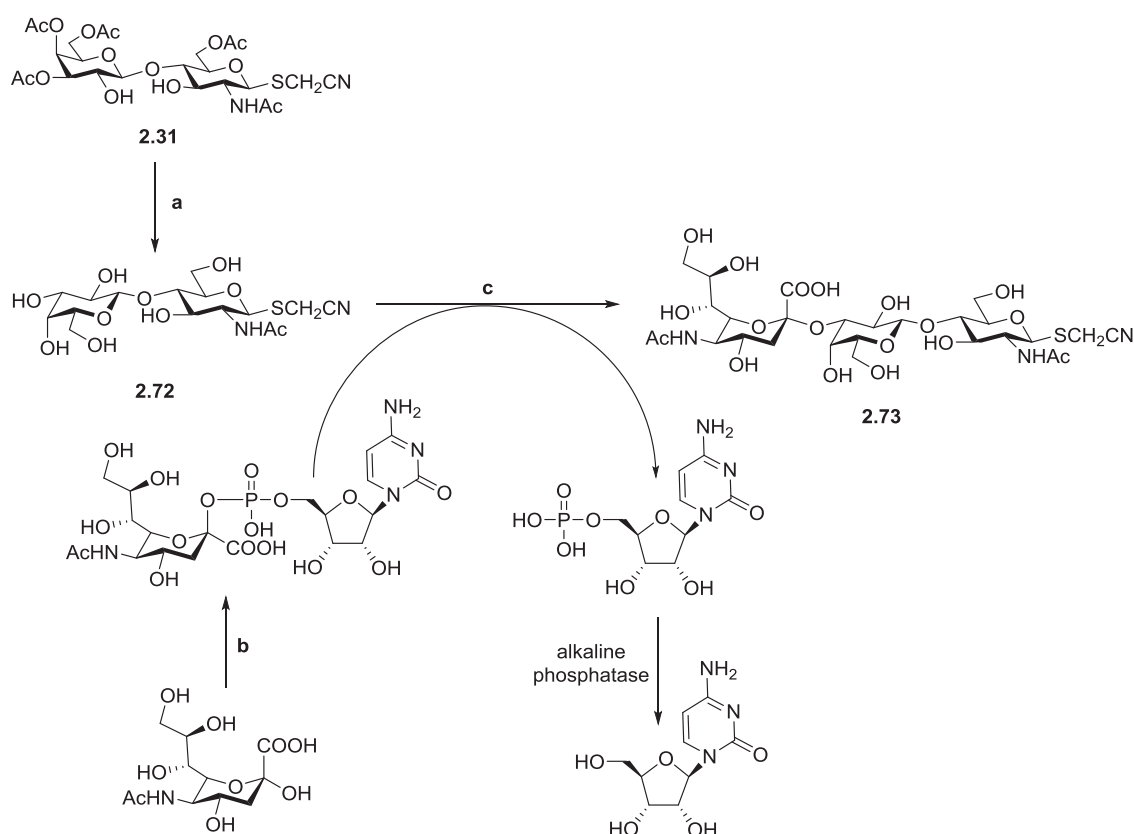
Le^a **2.29** was synthesised in 11 steps from three precursors (**2.6**, **2.10**, **2.20**) through two glycosylation reactions (Scheme 2.10).



Scheme 2.10 Reagents and conditions: a) DTBS-triflate, DMF, 4 °C, 77%; b) HBr (33% in AcOH), Ac₂O, r.t., 68%; c) DTBMP, AgOTf, DCM, MS, r.t., 84%; d) NEt₃·3HF, THF, r.t., quant.; e) AcCl, pyridine, DCM, -78 °C, 23%; f) 2.20, CuBr₂, TBAB, DCM:DMF (1:1), MS, r.t., quant.; g) i. CAN; ii. Ac₂O, pyridine, DMAP, ACN:H₂O (9:1), r.t., 54%.

Precursor **2.6** was prepared from GlcNAc **2.4** by reaction with DTBS-triflate at 4 °C, followed by quenching with pyridine. To form the disaccharidic core, donor **2.10** was reacted with acceptor **2.6** with the assistance of DTBMP and AgOTf in DCM. The reaction was run in the dark in dry conditions, and yielded the compound **2.25** in 77% yield after chromatographic purification. This disaccharide was then treated with NEt₃·3HF to deprotect the hydroxyls at positions 4 and 6 of the GlcNAc ring. The 6-OH was carefully re-protected by acetylation with AcCl and pyridine at very low temperature, to leave only the 5-OH free to react in the subsequent glycosylation. Fucosylation of **2.27** was performed using exactly the same protocol as for Le^x but reaction preceded in quantitative yield in this case. Removal of the PMBs and peracetylation were performed with the identical protocol used for Le^x. Final compound **2.29** was obtained in 54% yield.

sLacNAc was synthesised by a combination of chemical and enzymatic approaches. Previously synthesised (1,4)-LacNAc **2.31** was fully deprotected to the precursor **2.72** and sialylated using the protocol devised by Sugiarto et al.^[39] (Scheme 2.11).



Scheme 2.11 Reagents and conditions: a) **2.31**, NEt₃, MeOH, r.t., 94%; b) i. Neu, CTP, MgCl₂, NmCSS, r.t.; ii. **2.72**, Na₂CO₃, PmST, r.t., 76%.

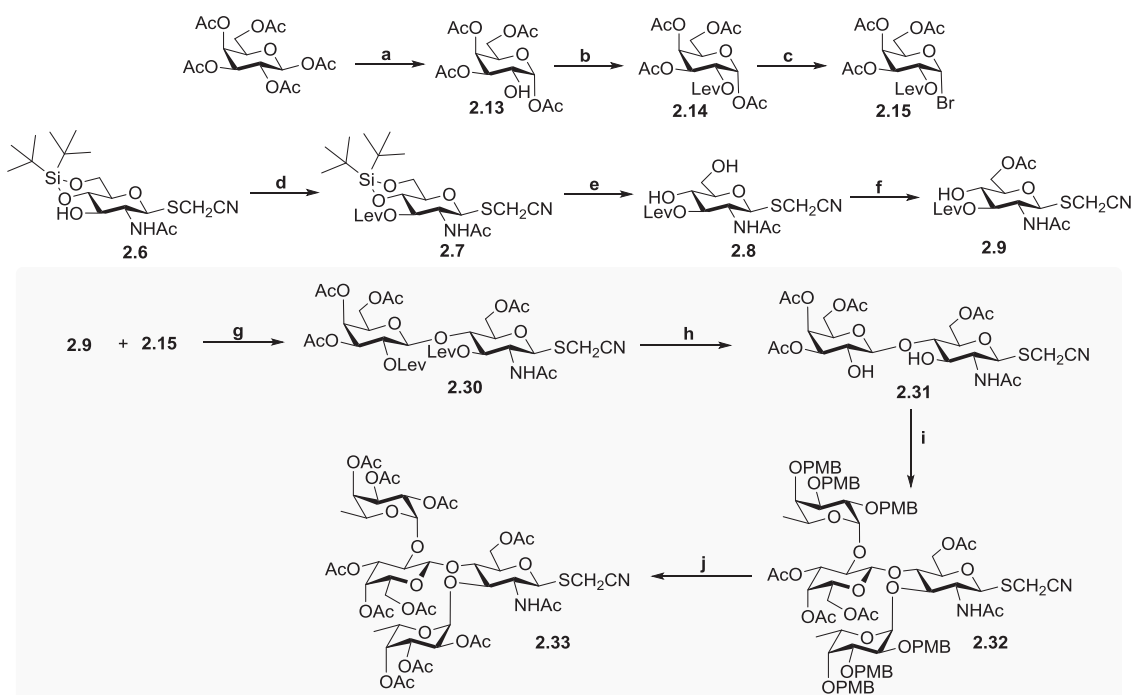
The disaccharide **2.31** was deprotected with triethylamine in MeOH for 48 h, to give the deprotected sugar **2.72** in 94% yield. Cytidine 5'-monophosphate-sialic acid (CMP-Neu) was formed *in situ* by addition of cytidine triphosphate (CTP), *Neisseria meningitidis* CMP-sialic acid synthetase (NmCSS)^[40] and MgCl₂

(cofactor of the enzyme) to neuraminic acid, yielding the substrate for the enzyme *Pasteurella multocida* sialyltransferase (PmST).^[41-43] PmST catalysed the transfer of the Neu from CMP-Neu to the 3-OH of the galactosyl residue of **2.72** to obtain the final product **2.73** in 76% yield. The CMP by-product is an inhibitor of PmST, therefore to allow the reaction to go to completion, we hydrolysed it to cytidine with alkaline phosphatase.

2.3.4 Tetrasaccharides

The synthesis of all tetrasaccharides in the Glycomap involves either chemical or enzymatic fucosylation, offering the opportunity to describe an example of a glycosylation reaction performed with these two different approaches.

Le^y **2.33** was synthesised in 12 steps from three precursors (**2.9**, **2.15**, **2.20**) through two glycosylation reactions (Scheme 2.12).

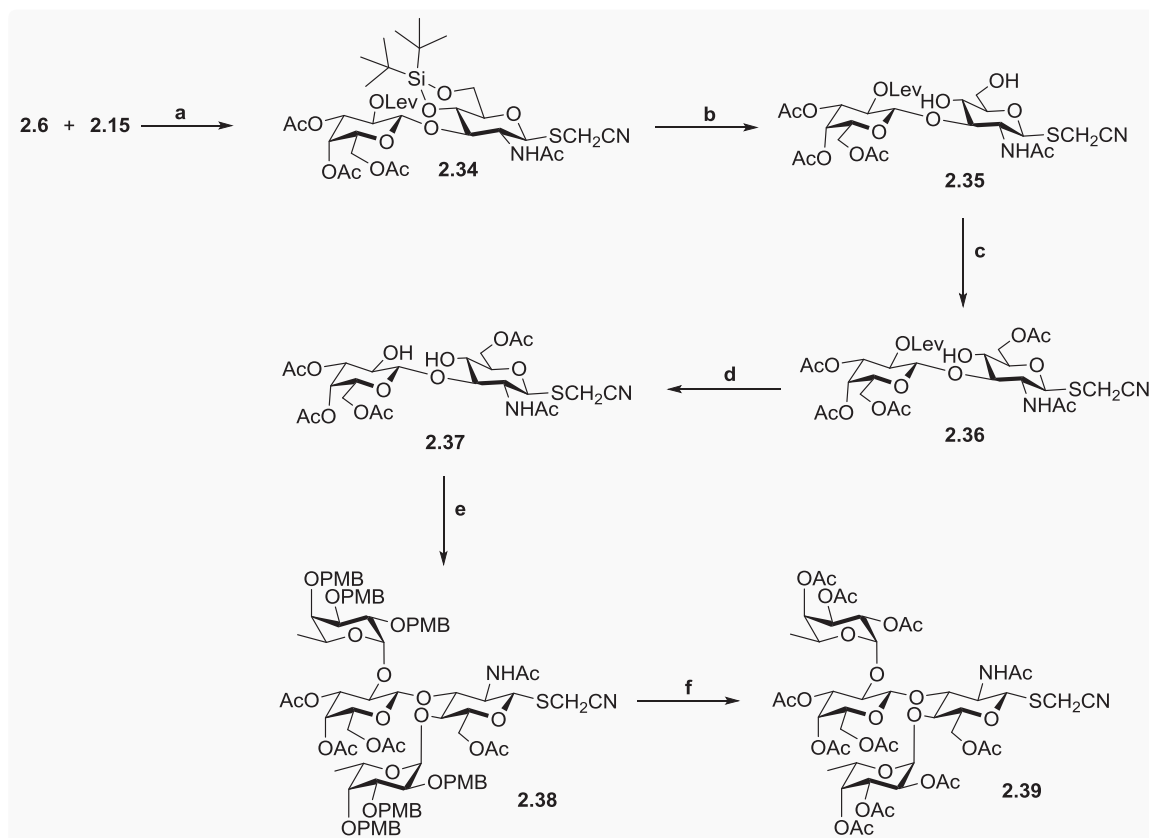


Scheme 2.12 Reagents and conditions: a) TFA, H₂O, r.t, quant.; b) EDAC·HCl, DMAP, levulinic acid, DCM, r.t., quant.; c) HBr (33% in AcOH), Ac₂O, r.t., quant.; d) EDAC·HCl, DMAP, levulinic acid, DCM, r.t., quant.; e) NEt₃·3HF, THF, 4 °C, quant.; f) AcCl, pyridine, -10 °C, 46%; g) AgOTf, DCM:toluene (1:1), -45 °C, 37%; h) hydrazine acetate, MeOH:DCM (1:1), r.t., 59%; i) 2.20, CuBr₂, TBAB, DCM:DMF (1:1), MS, r.t., 72%; j) i. CAN; ii. Ac₂O, pyridine, DMAP, ACN:H₂O (9:1), r.t., 71%.

Precursor **2.9** was prepared from the compound **2.6** by reacting the only free hydroxyl group with levulinic acid, removing the DTBS with NEt₃·3HF and carefully reprotecting the 6-OH with AcCl and pyridine at low temperature. Precursor **2.15** was obtained in three steps: the 2-OH was selectively deprotected in TFA (90% in water) and transformed into the levulinate derivative by reaction with levulinic acid; the sugar anomeric position was finally activated to the bromo-derivative by reacting with HBr and Ac₂O. The core disaccharide **2.30** was prepared from reaction of the donor **2.15** with acceptor **2.9** with the assistance of AgOTf at low temperature and in dry conditions (37% yield). After removal of the levulinic group with hydrazine acetate the second glycosylation took place. In strictly dry conditions, the fucosylation protocol used for the previous two Lewis

saccharides was applied here, to obtain a double addition on the two hydroxyl groups available in the disaccharide core in 72% yield. Once again PMBs of both Fuc residues were removed and substituted with acetyl groups by deprotection with CAN and peracetylation with Ac₂O to give the desired tetrasaccharide **2.33** in 71% yield.

Le^b **2.39** was synthesised in 11 steps from three precursors (**2.6**, **2.15**, **2.20**) through two glycosylation reactions (Scheme 2.13).

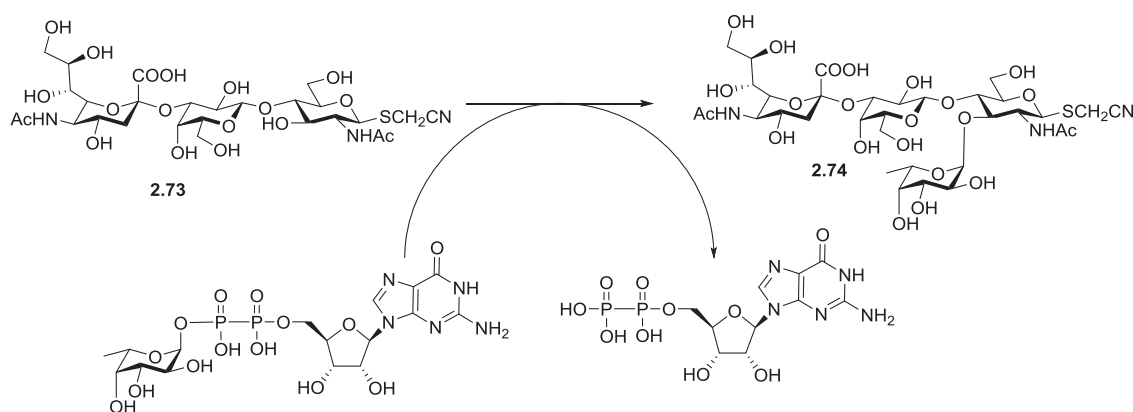


Scheme 2.13 Reagents and conditions: a) DTBMP, AgOTf, DCM, r.t., 63%; b) NEt₃·3HF, THF, r.t., 57%; c) AcCl, pyridine, DCM, -78 °C, 93%; d) hydrazine acetate, MeOH:DCM (1:1), r.t., 41%; e) **2.20**, CuBr₂, TBAB, DCM:DMF (1:1), MS, r.t., 40%; f) i. CAN; ii. Ac₂O, pyridine, DMAP, ACN:H₂O (9:1), r.t., 76%.

Donor **2.6** and acceptor **2.15** were reacted in dry conditions under activation using AgOTf and DTBMP to give the disaccharide **2.34** in 63% yield. This disaccharide

was deprotected by removal of the DTBS with $\text{NEt}_3 \cdot 3\text{HF}$, followed by re-protection of the 6-OH with AcCl and final removal of the levulinic acid with hydrazine acetate to yield the substrate **2.37** for the next glycosylation. Donor **2.20** reacted with both the free hydroxyl groups in a double addition to the disaccharide, following the same procedure used for **2.33**. Final tetrasaccharide **2.39** was obtained in 76% yield after the usual exchange of the PMB groups with acetyl groups.

sLe^x **2.74** was synthesised from sLacNAc **2.73** by enzymatic fucosylation similarly to the protocol described by Wang et al.^[44] (Scheme 2.14).



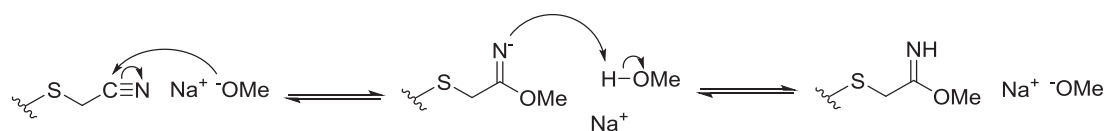
Scheme 2.14 Reagents and conditions: GDP-Fuc, MgCl_2 , FuT, 37 °C, quant.

Selective α -fucosylation is catalysed *in vivo* by fucosyltransferases (FuT) III,V or VII. Synthesis of the tetrasaccharide **2.74** was achieved enzymatically via FuT using GDP-Fuc as donor. Optimised condition for the reaction, adjusting the pH to 8 and reacting at 37 °C led to a complete conversion of **2.73** in the desired compound **2.74**. The reaction was scaled up to 100 mg with no decrease in yield.

As a final remark, we would like to emphasise how enzymatic glycosylations represent (at least in the synthesis we performed) the only glycosylations where no protecting groups were used and the product was isolated without the need for chromatographic purification.

2.4 2-Imino-2-methoxyethyl (IME) reagents

Prior to coupling on different macromolecular systems the sugar library of our Glycomap needs to be activated. The SCM group, which has been efficiently 'masking' the IME functionality, was transformed into the reactive derivative using NaOMe as activating reagent. NaOMe proved to be more efficient than other reagents^[45] and in our case it is particularly convenient as at the same time it deprotects the sugars from the acetyl groups. The deprotection occurs in a matter of minutes, regardless of the complexity of the saccharide, while the activation to IME requires hours. The reason being that the IME formation is an equilibrium reaction (Scheme 2.15). Rigorous dry conditions have no major impact on the formation of the IME, nonetheless reasonably dry conditions limit the hydrolysis of the IME into the carboxylic acid by-product.



Scheme 2.15 Mechanism of NaOMe-assisted IME formation. Reagents and conditions: NaOMe, MeOH, r.t., ~50%.

After reaction with NaOMe we always ended up with a mixture of deprotected sugar-SCM and deprotected sugar-IME. The ratio between the two has been determined by NMR (Fig. 2.6) from the ratio of the signal recorded for H-1 of both sugars in the mixture.

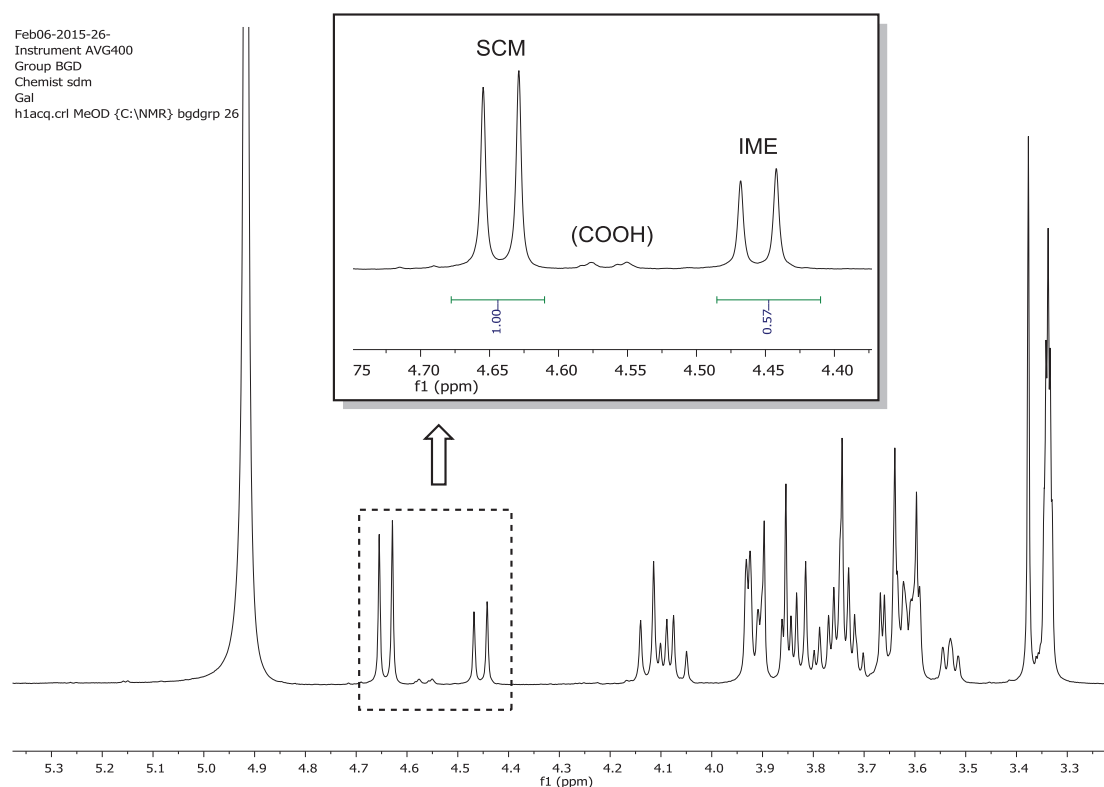


Figure 2.6 Determination of IME conversion by ^1H -NMR. Highlighted, the peaks for the H-1 protons of the sugars after the activation with NaOMe. From the ratio between their integrals it is possible to calculate the IME conversion.

Neu **2.70** and GlcA **2.66** carboxylic groups were deprotected with KOH 0.2 M before activation with NaOMe.

IME reacts with primary amines forming an amidine, which mimics the amide bonds and as a consequence it is more stable against hydrolysis and protease activity *in vivo*.

2.5 Conclusions

In this chapter we introduced the sugars we selected for the Glycomap and described their synthesis up to the reagents that will be coupled on different macromolecular systems in the next chapters. The sugars have been selected following a logical chemical evolution from β -D-Glc, to other monosaccharides, to di-, tri- and tetra-saccharides. In our selection we could not possibly include all the chemical modifications and ramifications available, so we sacrificed some interesting variables to limit the library to an achievable number for study on different macromolecular systems and offer a selected representation of their potential.

In the process of synthesising this sugar library we encountered a variety of the challenges that carbohydrate chemists regularly face. From the most complex design of a synthetic route to the tortuous chromatographic purifications (and the way around them). We cannot reiterate enough how every sugar is singular in its nature, and how minimal differences in their structure completely affect their chemical-physical properties and reactivity. We had the opportunity to experience the practical and theoretical principles of sugar reactivity through their modifications and conjugations.

We anticipate that expanding this Glycomap by exploring different chemical modifications and amplifications, different support systems, linkage-type, biological involvement (both in pathological and non-pathological states), in different animal models and more, will become of great interest to the scientific community in years to come.

2.6 References

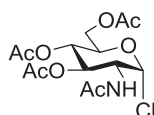
- [1] S. J. Angyal, *Angew. Chemie (International ed.)* **1969**, 8, 157-66.
- [2] S. Kornfield, E. Li, and I. Tabas, *J. Biol. Chem.* **1978**, 253, 7771-8.
- [3] D. J. Harvey, A. H. Merry, L. Royle, M. P. Campbell, R. A. Dwek, and P. M. Rudd, *Proteomics* **2009**, 9, 3796-3801.
- [4] A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, J. D. Marth, C. R. Bertozzi, G. W. Hart, and M. E. Etzler, *Proteomics* **2009**, 9, 5398-9.
- [5] D. J. Harvey, A. H. Merry, L. Royle, M. P. Campbell, and P. M. Rudd, *Proteomics* **2011**, 11, 4291-5.
- [6] T. J. Boltje, T. Buskas, and G-J. Boons, *Nat. Chem.* **2009**, 1, 611-622.
- [7] H. M. Christensen, S. Oscarson, and H. H. Jensen, *Carbohydr. Res.* **2015**, 408, 51-95.
- [8] J. Guo and X-S. Ye, *Molecules* **2010**, 15, 7235-65.
- [9] D. R. Mootoo, P. Konradsson, U. Udodong, and B. Fraser-Reid, *J. Am. Chem. Soc.* **1988**, 110, 5583-4.
- [10] N. L. Douglas, S. V. Ley, U. Lücking, and S. L. Warriner, *J. Chem. Soc. Perkin Trans. 1* **1988**, 51-66.
- [11] Z. Zhang, I. R. Ollmann, X-S. Ye, R. Wischnat, T. Baasov, and C-H. Wong, *J. Am. Chem. Soc.* **1999**, 121, 734-53.
- [12] X. Huang, L. Huang, H. Wang, and X- S. Ye, *Angew. Chemie Int. Ed.* **2004**, 43, 5221-24.
- [13] H. M. I. Osborn and T. H. Khan, *Tetrahedron* **1999**, 55, 1807-50.

- [14] P. H. Seeberger and W. C. Haase, *Chem. Rev.* **2000**, 100, 4349-93.
- [15] P. H. Seeberger, *Chem. Soc. Rev.* **2008**, 37, 19-28.
- [16] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, 85, 2149-54.
- [17] M. Marradi, D. Alcántara, J. M. de la Fuente, M. L. García-Martín, S. Cerdán, and S. Penadés, *Chem. Commun. (Camb)*. **2009**, 3922-4.
- [18] I. T. Horváth and J. Rábai, *Science* **1994**, 266, 72-75.
- [19] F. Zhang, W. Zhang, Y. Zhang, and D. P. Curran, *J. Org. Chem.* **2009**, 2594-97.
- [20] F. A. Jaipuri and N. L. Pohl, *Org. Biomol. Chem.* **2008**, 6, 2686-91.
- [21] G. S. Chen and N. L. Pohl, *Org. Lett.* **2008**, 10, 785-8.
- [22] X. He and T. Chan, *Synthesis* **2006**, 2006, 1645-51.
- [23] A. K. Pathak, C. K. Yerneni, Z. Young, and V. Pathak, *Org. Lett.* **2008**, 10, 145-8.
- [24] C. K. Yerneni, V. Pathak, and A. K. Pathak, *J. Org. Chem.* **2009**, 74, 6307-10.
- [25] J-Y. Huang, M. Lei, and Y-G. Wang, *Tetrahedron Lett.* **2006**, 47, 3047-50.
- [26] G. Nagy, N. L. B. Pohl, *J. Amer. Soc. Mass Spectrom.* **2015**, 26, 677-685
- [27] K. Bock and C. Pedersen, *J. Chem. Soc. Perkin Trans. 2* **1974**, 293.
- [28] D. R. Boyd, C. G. Watson, E. D. Becker, R. B. Bradley, and D. M. Jerina, *J. Am. Chem. Soc.* **1972**, 94, 8501-04.
- [29] V. M. S. Gil and A. C. P. Alves, *Mol. Phys.* **1969**, 16, 527-8.
- [30] V. M. S. Gil and J. J. C. Teixeira-Dias, *Mol. Phys.* **1968**, 15, 47-55.
- [31] T. Yonezawa, *J. Mol. Spect.* **1969**, 315, 341-5.
- [32] T. Yonezawa and I. Morishima, *J. Mol. Spect.* **1968**, 217, 210-7.

- [33] J. Albrand, A. Cogne, D. Gagnaire, and J. Robert, *Tetrahedron* **1971**, 27, 2453-61.
- [34] R. M. Schmaltz, S. R. Hanson, and C. H. Wong, *Chem. Rev.* **2011**, 111, 4259-4307.
- [35] D. B. Werz and P. H. Seeberger, *Chem. Eur. J.* **2005**, 11, 3194-3206.
- [36] B. G. Davis, *Chem. Ind.* **2000**, 134-8.
- [37] S. I. van Kasteren, S. J. Campbell, S. Serres, D. C. Anthony, N. R. Sibson, and B. G. Davis, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, 106, 18-23.
- [38] Saliba, R.C. DPhil thesis, University of Oxford: Oxford, **2013**.
- [39] G. Sugiarto, K. Lau, J. Qu, Y. Li, S. Lim, S. Mu, J. B. Ames, A. J. Fisher, and X. Chen, *ACS Chem. Biol.* **2012**, 7, 1232-40.
- [40] H. Yu, H. Yu, R. Karpel, and X. Chen, *Bioorg. Med. Chem.* **2004**, 12, 6427-35.
- [41] H. Yu, H. Chokhawala, R. Karpel, H. Yu, B. Wu, J. Zhang, Y. Zhang, Q. Jia, and X. Chen, *J. Am. Chem. Soc.* **2005**, 127, 17618-9.
- [42] L. Ni, M. Sun, H. Yu, H. Chokhawala, X. Chen, and A. J. Fisher, *Biochemistry* **2006**, 45, 2139-48.
- [43] L. Ni, H. A. Chokhawala, H. Cao, R. Henning, L. Ng, S. Huang, H. Yu, X. Chen, and A. J. Fisher, *Biochemistry* **2007**, 46, 6288-98.
- [44] W. Wang, T. Hu, P. A. Frantom, T. Zheng, B. Gerwe, D. S. Del Amo, S. Garret, R. D. Seidel, and P. Wu, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, 106, 16096-101.
- [45] Y. C. Lee, C. P. Stowell, and M. J. Krantz, *Biochemistry* **1976**, 15, 3956-63.

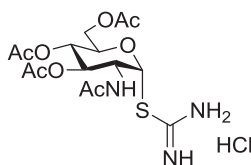
2.7 Supporting information

2-N-acetamido-3,4,6-tri-O-acetyl-2-deoxy- α -D-glucopyranosyl chloride **2.1**^[1]



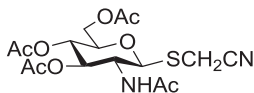
To acetyl chloride (58.1 mL, 817.1 mmol, 9 eq.) was added N-acetyl glucosamine (20 g, 90.4 mmol, 1 eq.) under nitrogen atmosphere and strong stirring. The suspension was stirred for 2 days at r.t., before CHCl_3 (200 mL) was added and the mixture was poured into ice (200 g) and H_2O (50 mL). The organic layer was separated and poured into a satd. solution of NaHCO_3 (250 mL) and ice (100 g). The mixture was stirred and then shaken into a separatory funnel until the production of gas ceased. The organic phase was then dried over MgSO_4 and evaporated to a small volume under vacuum. Et_2O was rapidly added and the product left to crystallise at r.t.. The mother liquor was evaporated and re-crystallised again from Et_2O . The product **2.1** was obtained as a white solid (29.07 g, 79.5 mmol, 88%). **TLC** R_f = 0.9 (EtOAc, 100%); **^1H NMR** (CDCl_3 , 400 MHz) δ 6.18 (d, J = 3.7 Hz, 1 H, H-1), 5.83 (d, J = 8.7 Hz, 1 H, NH), 5.32 (t, J = 11.9 Hz, 1 H, H-3), 5.21 (t, J = 9.8 Hz, 1 H, H-4), 4.53 (ddd, J = 10.7, 8.8, 3.7 Hz, 1 H, H-2), 4.30-4.24 (m, 2 H, H-5, H-6a), 4.13 (dd, J = 12.0, 1.6 Hz, 1 H, H-6b), 2.10, 2.05, 1.98 (3 x s, 12 H, 4 x Ac); **^{13}C NMR** (CDCl_3 , 101 MHz) δ 171.49, 170.58, 170.09, 169.13 (4 x CO), 93.64 (C-1), 70.91 (C-5), 70.14 (C-3), 66.96 (C-4), 61.15 (C-6), 53.51 (C-2), 23.09 (NHAc), 20.71, 20.68, 20.55 (3 x OAc); **MS** (ESI^+) m/z = 366.76 $[\text{M}+\text{H}]$; $[\alpha]_D^{20}$ = +118.9 (c = 1.04, CHCl_3);^[2] **IR** (ATR) 1746.09 (COOR), 1667.69 (CONH).

2-N-acetamido-3,4,6-tri-O-acetyl-2-deoxy- α -D-glucopyranosyl thiouronium chloride **2.2^[3]**



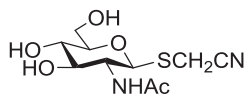
To a solution of **2.1** (29.07 g, 79.5 mmol, 1 eq.) in acetone (200 mL) was added thiourea (12.1 g, 159 mmol, 2 eq.). The mixture was stirred for 1 h at reflux (80 °C). The reaction was cooled to r.t., filtered and the solid washed with cold EtOH, to afford the title compound **2.2** as a white solid (32.85 g, 74.3 mmol, 93%). **TLC** R_f = 0.0 (EtOAc, 100%); **¹H NMR** (DMSO- d_6 , 400 MHz) δ 9.41, 9.21 (2 x s, 4 H, NH₂, NH₂Cl), 8.43 (d, J = 9.3 Hz, 1 H, NH), 5.68 (d, J = 10.5 Hz, 1 H, H-1), 5.13 (t, J = 9.8 Hz, 1 H, H-4), 4.93 (t, J = 9.7 Hz, 1 H, H-3), 4.24-4.15 (m, 2 H, H-5, H-6a), 4.07-3.97 (m, 2 H, H-2, H-6b), 2.01, 1.98, 1.93, 1.80 (4 x s, 12 H, 4 x Ac); **¹³C NMR** (DMSO- d_6 , 101 MHz) δ 170.50, 170.36, 170.06, 169.73 (4 x CO), 167.69 (CNH₂NH₂), 81.10 (C-1), 75.18 (C-5), 73.13 (C-3), 68.31 (C-4), 61.93 (C-6), 51.69 (C-2), 22.97 (NHAc), 21.01, 20.85, 20.74 (3 x OAc); **MS** (ESI⁺) m/z = 406.42 [M+H]⁺; $[\alpha]_D^{20}$ = -28.0 (c = 1.04, H₂O); **IR** (ATR) 3219.6, 3185.60 (NH₂), 1759.10 (COOR), 1646.90 (CONH).

2-N-acetamido-3,4,6-tri-O-acetyl-2-deoxy-1-thio-(S-cyanomethyl)- β -D-glucopyranoside **2.3^[4]**



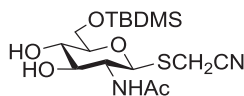
To a solution of **2.2** (20 g, 49.3 mmol, 1 eq.) in H₂O:acetone (1:1, 220 mL) were added Na₂S₂O₅ (20.6 g, 108.5 mmol, 2.2 eq.), K₂CO₃ (8.86 g, 64.1 mmol, 1.3 eq.) and chloroacetonitrile (59.3 mL, 937 mmol, 19 eq.). The mixture was stirred for 16 h at r.t., then it was poured into cold H₂O (400 mL) and stirred for further 1.5 h at r.t.. The reaction was extracted with CHCl₃ (3 x 250 mL), the organic layer backwashed with brine (200 mL), dried over MgSO₄ and evaporated under vacuum, to afford the title compound **2.3** as a white solid (17.21 g, 42.8 mmol, 87%). **TLC** R_f = 0.8 (EtOAc:MeOH, 4:1); **¹H NMR** (CDCl₃, 250 MHz) δ 5.96 (d, *J* = 9.2 Hz, 1 H, NH), 5.20 (t, *J* = 9.4 Hz, 1 H, H-3), 5.09 (t, *J* = 9.4 Hz, 1 H, H-4), 4.79 (d, *J* = 10.4 Hz, 1 H, H-1), 4.16-4.27 (m, 3 H, H-2, H-6), 3.77 (ddd, *J* = 9.2, 4.3, 2.4 Hz, 1 H, H-5), 3.32, 3.67 (2 x d, *J* = 17.0 Hz, 2 H, CH₂CN), 2.09, 2.04, 1.96 (3 x s, 12 H, 4 x Ac); **¹³C NMR** (CDCl₃, 63 MHz) δ 171.63, 171.17, 171.03 (3 x COO), 169.75 (CONH), 116.67 (CN), 83.44 (C-1), 76.66 (C-5), 73.69 (C-3), 68.49 (C-4), 62.30 (C-6), 53.19 (C-2), 23.58 (NHAc), 21.20, 21.11, 21.05 (3 x OAc), 15.11 (CH₂CN); **MS** (ESI⁺) *m/z* = 403.42 [M+H]; $[\alpha]_D^{20}$ = -94.3 (c = 1.0, CHCl₃);^[4] **IR** (ATR) 2198.15 (CN), 1738.74 (COOR), 1660.67 (CONH).

2-N-acetamido-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.4**^[4]



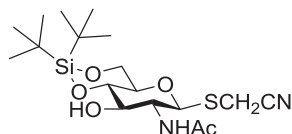
To a solution of **2.3** (23 g, 57 mmol, 1 eq.) in MeOH (445 mL) was added NEt₃ (44.5 mL, 319 mmol, 5.6 eq.). The mixture was stirred for 16 h at r.t., then the solvent was removed under vacuum and the slurry obtained co-evaporated twice with toluene and once with Et₂O, to afford the title compound **2.4** as a white solid (19 g, 57 mmol, q.). **TLC** R_f = 0.2 (EtOAc:MeOH, 4:1); **¹H NMR** (DMSO, 400 MHz) δ 7.89 (d, J = 9.2 Hz, 1 H, NH), 5.11 (dd, J = 9.0, 5.4 Hz, 2 H, OH-3, OH-4), 4.50 (dd, J = 11.3, 4.3 Hz, 2 H, H-1, OH-6), 3.79 (dd, J = 41.5, 17.0 Hz, 2 H, CH₂CN), 3.74 - 3.65 (m, 1 H, H-6a), 3.56 (dd, J = 19.5, 9.8 Hz, 1 H, H-2), 3.45 (dt, J = 11.7, 5.7 Hz, 1 H, H-6b), 3.34 - 3.29 (m, 1 H, H-3), 3.12 (ddd, J = 13.7, 9.0, 3.5 Hz, 2 H, H-4, H-5); **¹³C NMR** (DMSO, 126 MHz) δ 169.31 (CO), 118.10 (CN), 83.11 (C-1), 81.44 (C-5), 74.97 (C-3), 70.45 (C-4), 61.29 (C-6), 54.07 (C-2), 22.90 (Ac), 13.79 (CH₂CN); **MS** (ESI⁺) m/z = 277.32 [M+H]; $[\alpha]_D^{20}$ = -64.6 (c = 1.0, H₂O);^[4] **IR** (ATR) 3428.88 (OH), 2196.43 (CN), 1664.21 (CONH).

2-N-acetamido-6-O-*tert*-butyl-dimethylsilyl-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.5**^[5]



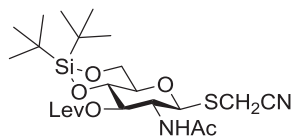
To a solution of **2.4** (8.2 g, 29.7 mmol, 1 eq.) in DMF (60 mL) at 0 °C, were added imidazole (3.03 g, 44.5 mmol, 1.5 eq.) and TBDMSCl (4.92 g, 32.6 mmol, 1.1 eq.). The mixture was stirred for 16 h at r.t., then the solvent was removed under vacuum and the oil obtained precipitated with H₂O. The precipitate was filtered, washed with H₂O and co-evaporated with toluene and acetone, to afford the title compound **2.5** as a white solid (7.72 g, 19.8 mmol, 67%). **TLC** R_f = 0.85 (EtOAc:MeOH, 9:1); **¹H NMR** (acetone-d₆, 400 MHz) δ 7.32 (d, *J* = 9.0 Hz, 1 H, NH), 4.69 (d, *J* = 10.3 Hz, 1 H, H-1), 4.01 (dd, *J* = 11.3, 2.0 Hz, 1 H, H-6a), 3.89-3.78 (m, 2 H, H-2, H-6b), 3.69 (dd, *J* = 37.0, 17.6 Hz, 2 H, CH₂CH), 3.62 (dd, *J* = 10.5, 7.9 Hz, 1 H, H-3), 3.51-3.42 (m, 1 H, H-4), 3.35 (ddd, *J* = 9.7, 5.4, 2.0 Hz, 1 H, H-5), 1.89 (s, 3 H, Ac), 0.91 (s, 9 H, ^{*t*}Bu), 0.10, 0.09 (2 x s, 2 x 3 H, 2 x CH₃); **¹³C NMR** (acetone-d₆, 101 MHz) δ 169.74 (CO), 116.87 (CN), 83.29 (C-1), 81.15 (C-5), 76.10 (C-3), 70.65 (C-4), 62.90 (C-6), 54.39 (C-2), 25.41 (C(CH₃)₃), 22.25 (Ac), 18.10 (C(CH₃)₃), 13.61 (CH₂CN), -5.86, -5.96 (2 x CH₃); **MS** (ESI⁺) *m/z* = 391.58 [M+H]; [α]_D²⁰ = -89.8 (c = 1.0, acetone); **IR** (ATR) 3291.99 (OH), 2198.37 (CN), 1649.35 (CONH).

**2-N-acetamido-4,6-di-*tert*-butyl-dimethylsilylidene-1-thio-(S-cyanomethyl)-
2-deoxy- β -D-glucopyranoside **2.6****



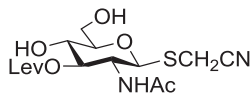
Into a solution of **2.4** (10 g, 36.2 mmol, 1 eq.) in DMF (dry, 160 mL) was bubbled argon. After about 30 min the reaction was cooled to $-40\text{ }^{\circ}\text{C}$, and it was slowly added to DTBS-triflate (12.38 mL, 38 mmol, 1.05 eq.). The mixture was stirred for 16 h at $4\text{ }^{\circ}\text{C}$, then pyridine (10 mL) was added and stirred for another 15 min. The solvent was removed under vacuum and the oil obtained was precipitated with H_2O . The precipitate was filtered, washed with H_2O and co-evaporated with MeOH, to afford the title compound **2.6** as a white solid (11.56 g, 27.7 mmol, 77%). **TLC** $R_f = 0.75$ (DCM:MeOH, 10:1); **^1H NMR** (MeOD, 400 MHz) δ 4.72 (d, $J = 10.4$ Hz, 1 H, H-1), 4.20 (dd, $J = 10.1, 5.1$ Hz, 1 H, H-6a), 3.95-3.80 (m, 2 H, H-2, H-6b), 3.79-3.64 (m, 2 H, CH_2aCN , H-4), 3.64-3.47 (m, 3 H, CH_2Bcn , H-3, H-5), 1.98 (s, 3 H, Ac), 1.08, 1.02 (2 x s, 2 x 9 H, 2 x ^tBu); **^{13}C NMR** (MeOD, 101 MHz) δ 172.53 (CO), 117.26 (CN), 84.48 (C-1), 77.85 (C-5), 75.41 (C-4), 75.02 (C-3), 66.26 (C-6), 54.55 (C-2), 26.54 (2 x $\text{C}(\text{CH}_3)_3$), 22.60 (2 x $\text{C}(\text{CH}_3)_3$), 21.77 (Ac), 14.09 (CH_2CN); **MS** (ESI $^+$) $m/z = 417.61$ [M+H]; $[\alpha]_D^{20} = -44.4$ ($c = 1.0$, acetone); **IR** (ATR) 3297.82 (OH), 2189.31 (CN), 1645.86 (CONH).

2-N-acetamido-4,6-di-*tert*-butyl-dimethylsilylidene-3-O-levunoyl-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.7**



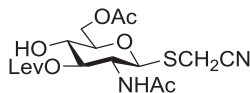
To **2.6** (11.56 g, 28.7 mmol, 1 eq.) were added EDAC hydrochloride (11.01 g, 57.4 mmol, 2 eq.) and DMAP (2.1 g, 17.23 mmol, 0.6 eq.). The mixture was dried under vacuum for 45 min, then DCM (100 mL) and LevOH (3.53 mL, 34.4 mmol, 1.2 eq.) were added under nitrogen atmosphere. The reaction was stirred for 2 h at r.t. and then diluted with DCM (750 mL), washed with H₂O (400 mL), HCl 1M (400 mL), satd. solution of NaHCO₃ (400 mL) and brine (400 mL). The organic layer was dried over MgSO₄ and evaporated to afford the title compound **2.7** as a white solid (14.81 g, 28.7 mmol, quant.). **TLC** R_f = 0.9 (DCM:MeOH, 10:1); **¹H NMR** (CDCl₃, 200 MHz) δ 6.09 (d, *J* = 9.2 Hz, 1 H, NH), 5.04 (dd, *J* = 10.2, 9.1 Hz, 1 H, H-3), 4.76 (d, *J* = 10.4 Hz, 1 H, H-1), 4.26-4.01 (m, 2 H, H-6a, H-2), 3.88 (t, *J* = 9.4 Hz, 2 H, H-6b, H-4), 3.57 (ddd, *J* = 14.8, 13.6, 5.3 Hz, 2 H, CH₂aCN, C-5), 3.29 (d, *J* = 17.1 Hz, 1 H, CH₂bCN), 2.83-2.38 (m, 4 H, Lev), 2.14 (s, 3 H, CH₃), 1.97 (s, 3 H, Ac), 1.00, 0.94 (2 x s, 2 x 9 H, 2 x ^tBu); **¹³C NMR** (CDCl₃, 50 MHz) δ 206.77 (COCH₃), 172.77 (CONHAc), 172.53 (COO), 116.88 (CN), 84.44 (C-1), 77.91 (C-5), 75.13 (C-4), 74.96 (C-3), 66.13 (C-6), 54.68 (C-2), 36.88 (CH₂COO), 30.01 (CH₂COCH₃), 28.14 (CH₃), 26.49 (2 x C(CH₃)₃), 22.11 (2 x C(CH₃)₃), 21.03 (Ac), 14.00 (CH₂CN); **MS** (ESI⁺) *m/z* = 515.71 [M+H]; [α]_D²⁰ = -77.9 (c = 1.0, MeOH); **IR** (ATR) 2189.31 (CN), 1734.52 (COO), 1719.68 (CO), 1664.48 (CONH).

2-N-acetamido-3-O-levunoyl-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside 2.8



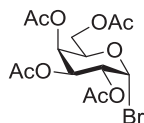
To a solution of **2.7** (14.81 g, 28.8 mmol, 1 eq.) in THF (dry, 300 mL) at 4 °C, was added NEt₃ trihydrofluoride (9 mL, 55.2 mmol, 2 eq.). The mixture was stirred for 16 h at 4 °C, the solvent removed under vacuum and the slurry obtained co-evaporated with toluene, to afford the title compound **2.8** as a yellow oil (10.81 g, 28.8 mmol, quant.). **TLC** R_f = 0.7 (DCM:MeOH, 10:1); **¹H NMR** (MeOD, 200 MHz) δ 5.06 (m, 1 H, H-3), 4.79 (d, *J* = 10.4 Hz, 1 H, H-1), 4.01-3.93 (m, 1 H, H-2), 3.92-3.86 (m, 1 H, H-6a), 3.83 (br, 1 H, SCH₂aCN), 3.74 (d, *J* = 5.4 Hz, 1 H, H-6b), 3.70-3.63 (m, 1 H, SCH₂bCN), 3.61-3.53 (m, 1 H, H-4), 3.49-3.40 (m, 1 H, H-5), 2.78 (dd, *J* = 12.9, 6.1 Hz, 2 H, CH₂COCH₃), 2.63-2.46 (m, 2 H, CH₂COO), 2.15 (s, 3 H, CH₃), 1.95 (s, 3 H, Ac); **¹³C NMR** (MeOD, 101 MHz) δ 207.64 (COCH₃), 172.67 (COO), 172.24 (CONH), 117.09 (CN), 82.72 (C-1), 80.86 (C-5), 76.31 (C-3), 68.18 (C-4), 61.14 (C-6), 52.47 (C-2), 37.14 (CH₂COCH₃), 28.27 (CH₂COO), 27.57 (CH₃), 21.42 (Ac), 13.42 (CH₂CN); **MS** (ESI⁺) *m/z* = 375.41 [M+H]; $[\alpha]_D^{20}$ = -57.7 (c = 1.0, MeOH); **IR** (ATR) 3289.31 (OH), 2187.16 (CN), 1739.95 (COO), 1716.56 (CO), 1660.42 (CONH).

2-N-acetamido-6-O-acetyl-3-O-levunoyl-1-thio-(S-cyanomethyl)- 2-deoxy- β -D-glucopyranoside 2.9



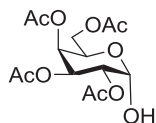
To a solution of **2.8** (17.81 g, 47.6 mmol, 1 eq.) in pyridine (dry, 240 mL) at -10 °C and under nitrogen, was slowly added AcCl (5 mL, 71.35 mmol, 1.5 eq.). The mixture was stirred for 2 h at -10 °C, then more AcCl was added (2.5 mL, 35.68 mmol, 0.75 eq.). After 9 h the reaction was quenched with MeOH (50 mL) and the solvent removed under vacuum. The crude was purified by flash column chromatography over silica (CHCl₃:EtOH, 10:0.25) to afford the title compound **2.9** as a white solid (8.87 g, mmol, 46%). **TLC** R_f = 0.75 (DCM:MeOH, 10:1); **¹H NMR** (MeOD, 400 MHz) δ 5.04 (dd, *J* = 10.1, 8.7 Hz, 1 H, H-3), 4.79 (d, *J* = 10.4 Hz, 1 H, H-1), 4.44 (dd, *J* = 12.0, 1.8 Hz, 1 H, H-6a), 4.25 (dd, *J* = 12.0, 5.5 Hz, 1 H, H-6b), 3.97 (t, *J* = 10.3 Hz, 1 H, H-2), 3.73 (d, *J* = 17.2 Hz, 1 H, SCH₂aCN), 3.68–3.53 (m, 3 H, SCH₂bCN, H-5, H-4), 2.90–2.68 (m, 2 H, CH₂COCH₃), 2.67–2.47 (m, 2 H, CH₂COO), 2.16 (s, 3 H, CH₃), 2.08 (s, 3 H, OAc), 1.95 (s, 3 H, NHAc); **¹³C NMR** (MeOD, 101 MHz) δ 207.59 (COCH₃), 172.63 (COO), 172.26 (CONH), 171.30 (CO-OAc), 116.73 (CN), 82.66 (C-1), 78.02 (C-5), 76.04 (C-3), 68.28 (C-4), 62.94 (C-6), 52.42 (C-2), 37.12 (CH₂COCH₃), 28.25 (CH₃), 27.54 (CH₂COO), 21.40 (NHAc), 19.34 (OAc), 13.56 (CH₂CN); **MS** (ESI⁺) *m/z* = 417.46 [M+H]; $[\alpha]_D^{20}$ = -77.8 (*c* = 1.0, MeOH); **IR** (ATR) 3410.48 (OH), 2176.81 (CN), 1736.62 (COO), 1718.44, 1711.53 (2 x CO), 1645.42 (CONH).

2,3,4,6-tetra-O-acetyl-1-bromo- α -D-galactopyranoside **2.10**



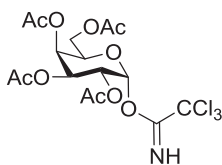
To 1,2,3,4,6-penta-O-acetyl- β -D-galactopyranoside (10.82 g, 27.75 mmol, 1 eq.) was added HBr (33% in AcOH, 40 mL) and acetic anhydride (10 mL). The mixture was stirred for 1 h at r.t.. After this time the reaction was poured into ice (200 g) and water (50 mL), recovered with EtOAc (200 mL) and separated. The organic layer was washed with H₂O (250 mL), NaHCO₃ satd. (250 mL), brine (250 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum to afford the title compound **2.10** as a yellow oil (7.71 g, 18.75 mmol, 68%). **TLC** R_f = 0.8 (EtOAc:PE, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 6.68 (d, J = 3.9 Hz, 1 H, H-1), 5.50 (d, J = 2.8 Hz, 1 H, H-4), 5.39 (dd, J = 10.6, 3.3 Hz, 1 H, H-3), 5.03 (dd, J = 10.6, 3.9 Hz, 1 H, H-2), 4.47 (t, J = 6.5 Hz, 1 H, H-5), 4.13 (ddd, J = 31.0, 11.4, 6.6 Hz, 1 H, H-6), 2.14, 2.10, 2.05, 2.00 (4 x s, 4 x 3H, 4 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.35, 170.09, 169.92, 169.79 (4 x CO), 88.13 (C-1), 71.06 (C-5), 67.99 (C-2), 67.76 (C-3), 66.97 (C-4), 60.84 (C-6), 20.77, 20.67, 20.61, 20.58 (4 x OAc); **MS** (ESI⁺) m/z = 412.21 [M+H]; $[\alpha]_D^{20}$ = -77.6 (c = 1.0, CDCl₃);^[6] **IR** (ATR) 2176.99 (CN), 1717.94 (CO).

2,3,4,6-tetra-O-acetyl- α -D-galactopyranoside **2.11**^[7]



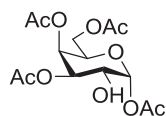
To a solution of β -D-pentaacetyl galactose (10 g, 25.6 mmol, 1 eq.) in DMF (dry, 100 mL) was added hydrazine acetate (2.6 g, 28.2 mmol, 1.1 eq.), under nitrogen. The mixture was stirred at 50 °C for 5 h, then recovered with EtOAc (200 mL), washed with satd. solution of NaHCO_3 (3 x 125 mL) and brine (125 mL). The organic layer was dried over MgSO_4 and evaporated to afford the title compound **2.11** as a colourless oil (9.87 g, 25.6 mmol, quant.). **TLC** R_f = 0.9 (EtOAc, 100%); **$^1\text{H NMR}$** (CDCl_3 , 200 MHz) δ 7.98 (s, 1 H, OH), 5.55-5.30 (m, 2 H, H-1, H-3), 5.10 (ddd, J = 8.3, 7.2, 2.8 Hz, 1 H, H-5), 4.55-4.21 (m, 2 H, H-2, H-4), 4.20-4.01 (m, 2 H, H-6), 2.13, 2.08, 2.03, 1.97 (4 x s, 4 x 3 H, 4 x OAc); **$^{13}\text{C NMR}$** (CDCl_3 , 50 MHz) δ 170.22 (CO), 94.87 (C-1), 74.74 (C-5), 71.73 (C-3), 69.99 (C-2), 69.38 (C-4), 62.70 (C-6), 21.40, 21.18, 20.75, 20.53 (4 x Ac); **MS** (ESI⁻) m/z = 347.29 [M-H]; $[\alpha]_D^{20}$ = -79.2 (c = 1.0, MeOH); **IR** (ATR) 3411.67 (OH), 1732.12 (CO).

2,3,4,6-tetra-O-acetyl- α -D-galactopyranoside trichloroacetimidate **2.12**^[8]



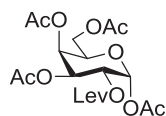
To a solution of **2.11** (9.87 g, 25.6 mmol, 1 eq.) in DCM (100 mL) were added trichloro acetonitrile (25.67 mL, 256 mmol, 10 eq.) and DBU (766 μ L, 5.12 mmol, 0.2 eq.), under nitrogen. The mixture was stirred at r.t. for 16 h, then evaporated and purified by flash column chromatography over silica (PE:EtOAc, 2:1 to 1:1) to afford the title compound **2.12** as a light yellow solid (9.15 g, 18.6 mmol, 73%). **TLC** R_f = 0.5 (PE:EtOAc, 2:1); **$^1\text{H NMR}$** (CDCl_3 , 400 MHz) δ 8.65 (s, 1 H, NH), 6.51 (d, J = 3.3 Hz, 1 H, H-1), 5.47 (d, J = 2.0 Hz, 1 H, H-4), 5.30 (ddd, J = 26.2, 10.9, 3.2 Hz, 2 H, H-3, H-2), 4.36 (t, J = 6.5 Hz, 1 H, H-5), 4.08 (dd, J = 11.3, 6.5 Hz, 1 H, H-6a), 4.04-3.88 (m, 1 H, H-6b), 2.09, 1.95, 1.93, 1.93 (4 x s, 4 x 3 H, 4 x OAc); **$^{13}\text{C NMR}$** (CDCl_3 , 101 MHz) δ 170.18, 170.03, 169.97, 169.87 (4 x CO), 160.74 (CNH), 93.43 (C-1), 90.70 (CCl_3), 68.93 (C-5), 67.42, 67.30 (C-3, C-4), 66.82 (C-2), 61.18 (C-6), 20.58, 20.54, 20.53, 20.46 (4 x OAc); **MS** (ESI^+) m/z = 493.70 [$\text{M}+\text{H}$]; $[\alpha]_D^{20}$ = +105.78 (c = 1.0, CDCl_3);^[9] **IR** (ATR) 1745.24 (CO).

1,3,4,6-tetra-O-acetyl- α -D-galactopyranoside **2.13**



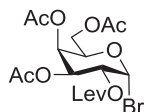
To β -D-pentaacetyl galactose (10 g, 25.6 mmol, 1 eq.) was added TFA (90% in H_2O , 40 mL). The mixture was stirred at r.t for 6.5 h, then evaporated to afford the title compound **2.13** as colourless oil (9.23 g, 25.6 mmol, quant.). **TLC** R_f = 0.9 (EtOAc, 100%); **$^1\text{H NMR}$** (CDCl_3 , 200 MHz) δ 6.30 (d, J = 3.9 Hz, 1 H, H-1), 5.56-5.34 (m, 1 H, H-4), 5.19 (dd, J = 10.6, 3.3 Hz, 1 H, H-3), 4.35-3.87 (m, 4 H, H-5, H-2, H-6), 2.19, 2.15, 2.07, 2.05 (4 x s, 4 x 3 H, 4 x OAc), 1.37 (d, J = 6.3 Hz, 1 H, OH); **$^{13}\text{C NMR}$** (CDCl_3 , 63 MHz) δ 171.26, 170.88, 170.53, 169.87 (4 x CO), 92.33 (C-1), 70.81 (C-5), 69.11 (C-3), 67.92 (C-4), 66.47 (C-6), 61.70 (C-2), 21.05, 21.03, 21.02, 21.00 (4 x Ac); **MS** (ESI^-) m/z = 347.29 $[\text{M}-\text{H}]^-$; $[\alpha]_D^{20}$ = +148.00 (c = 1.0, CDCl_3);^[10] **IR** (ATR) 3357.92 (OH), 1745.54 (CO).

1,3,4,6-tetra-O-acetyl-2-O-levulinoyl- α -D-galactopyranoside **2.14**^[11]



To a mixture of **2.13** (9.23 g, 25.6 mmol, 1 eq.), EDAC hydrochloride (9.8 g, 51.2 mmol, 2 eq.) and DMAP (1.88 g, 15.36 mmol, 0.6 eq.), dried under vacuum for 30 min, were added DCM (100 mL) and levulinic acid (3.15 mL, 30.7 mmol, 1.2 eq.), under nitrogen. The mixture was stirred at r.t. for 2 h, then diluted with DCM (150 mL), washed with H₂O (125 mL), HCl 1M (125 mL), satd. solution of NaHCO₃ (125 mL) and brine (125 mL). The organic layer was dried over MgSO₄ and evaporated to afford the title compound **2.14** as a yellow oil (11.83 g, 25.6 mmol, quant.). **TLC** R_f = 0.4 (EtOAc:PE, 1:1); **¹H NMR** (CDCl₃, 200 MHz) δ 6.35 (d, J = 2.9 Hz, 1 H, H-1), 5.55-5.44 (m, 1 H, H-4), 5.37-5.29 (m, 2 H, H-3, H-2), 4.33 (td, J = 6.7, 1.0 Hz, 1 H, H-5), 4.17-3.94 (m, 2 H, H-6), 2.88-2.36 (m, 4 H, Lev), 2.17, 2.16, 2.15, 2.12, 2.03 (5 x s, 5 x 3 H, 4 x OAc, CH₃); **¹³C NMR** (CDCl₃, 63 MHz) δ 206.42 (COCH₃), 172.10 (COO), 170.74, 170.63, 170.50, 169.38 (4 x CO-OAc), 89.96 (C-1), 69.12 (C-5), 67.84 (C-4), 67.56 (C-3), 67.03 (C-6), 61.63 (C-2), 38.03 (CH₂COO), 30.08 (CH₂COCH₃), 27.99 (CH₃), 21.28, 21.17, 21.11, 21.01 (4 x Ac); **MS** (ESI⁺) m/z = 447.41 [M+H]⁺; $[\alpha]_D^{20}$ = +86.60 (c = 1.0, CDCl₃);^[12] **IR** (ATR) 1743.87 (CO).

1-bromo-3,4,6-tri-O-acetyl-2-O-levulinoyl- α -D-galactopyranoside **2.15**^[13]



To a solution of **2.14** (11.82 g, 25.6 mmol, 1 eq.) in Ac₂O (4 mL, 42.3 mmol, 1.6 eq.), was added HBr (33% in AcOH, 26 mL, 145 mmol, 5.6 eq.). The mixture was stirred at r.t for 3 h, then diluted with EtOAc (200 mL), washed with H₂O (125 mL), satd. solution of NaHCO₃ (3 x 125 mL) and brine (125 mL). The organic layer was dried over MgSO₄ and evaporated to afford the title compound **2.15** as a yellow oil (11.7 g, 25.6 mmol, quant.). **TLC** R_f = 0.35 (EtOAc:PE, 1:1); **¹H NMR** (CDCl₃, 200 MHz) δ 6.64 (d, *J* = 4.0 Hz, 1 H, H-1), 5.59-5.34 (m, 3 H, H-4, H-3), 5.06 (ddd, *J* = 10.7, 7.9, 3.9 Hz, 1 H, H-2), 4.48 (t, *J* = 6.5 Hz, 1 H, H-5), 4.26-3.97 (m, 2 H, H-6), 2.89-2.52 (m, 4 H, Lev), 2.18, 2.15, 2.06, 2.04 (4 x s, 4 x 3 H, 3 x OAc, CH₃); **¹³C NMR** (CDCl₃, 50 MHz) δ 205.66 (COCH₃), 171.97 (COO), 170.56, 170.34, 169.87 (3 x CO-Ac), 88.65 (C-1), 69.99 (C-5), 67.97 (C-4), 66.81 (C-3), 65.92 (C-6), 61.79 (C-2), 37.13 (CH₂COO), 32.18 (CH₂COCH₃), 29.01 (CH₃), 21.33, 21.14, 21.08 (3 x Ac); **MS** (ESI⁻) *m/z* = 466.25 [M-H]; [α]_D²⁰ = +186.80 (*c* = 1.0, CDCl₃); **IR** (ATR) 1747.69 (CO).

1,2,3,4-tetra-O-acetyl-6-deoxy- β -L-galactopyranoside **2.16**^[14]



To a suspension of L-fucose (15 g, 91.3 mmol, 1 eq.) DCM (175 mL), were added DMAP (1.1 g, 9.13 mmol, 0.1 eq.) and triethylamine (102 mL, 731.8 mmol, 8 eq.). The mixture was stirred at 0 °C and Ac₂O (52 mL, 550 mmol, 6 eq.) was added. The reaction was then stirred at r.t for 16 h, poured into ice/water (200 mL), the organic layer separated and washed with water (125 mL), HCl 1M (2 x 125 mL), satd. solution of NaHCO₃ (2 x 125 mL) and brine (125 mL). The organic layer was dried over MgSO₄ and evaporated to afford the title compound **2.16** as an orange oil (30 g, 91.3 mmol, quant.). **TLC** R_f = 0.9 (EtOAc:PE, 4:1); **¹H NMR** (CDCl₃, 200 MHz) δ 6.33 (d, J = 1.8 Hz, 1 H, H-1), 5.39-5.20 (m, 3 H, H-2, H-3, H-4), 4.26 (q, J = 6.5 Hz, 1 H, H-5), 2.22, 2.17, 2.14, 2.01 (4 x s, 4 x 3 H, 4 x OAc), 1.15 (d, J = 6.5 Hz, 3 H, H-6); **¹³C NMR** (CDCl₃, 63 MHz) δ 170.93, 170.58, 170.36, 169.87 (4 x CO), 90.31 (C-1), 71.62 (C-5), 70.95 (C-3), 70.32 (C-4), 67.65 (C-2), 21.30, 21.23, 21.06, 20.96 (4 x Ac), 16.30 (C-6); **MS** (ESI⁺) m/z = 333.31 [M+H]; $[\alpha]_D^{20}$ = +6.87 (c = 1.0, CDCl₃); **IR** (ATR) 1752.34 (CO).

1-bromo-2,3,4-tri-O-acetyl-6-deoxy- β -L-galactopyranoside **2.17**^[15]



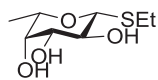
To a solution of **2.16** (33.68 g, 91.3 mmol, 1 eq.) in DCM (dry, 40 mL), were added HBr (33% in AcOH, 39.6 mL, 228.25 mmol, 2.5 eq.) and Ac₂O (6.9 mL, 73.04 mmol, 0.8 eq.). The mixture was stirred at r.t. for 1 h, poured into ice/water (250 mL), recovered with EtOAc (200 mL) and separated. The organic layer was washed with H₂O (125 mL), satd. solution of NaHCO₃ (3 x 125 mL), dried over MgSO₄ and evaporated to afford the title compound **2.17** as a yellow oil (31.25 g, 88.5 mmol, 97%). **TLC** R_f = 0.85 (EtOAc:PE, 4:1); **¹H NMR** (CDCl₃, 200 MHz) δ 6.69 (d, J = 4.0 Hz, 1 H, H-1), 5.53-5.15 (m, 2 H, H-2, H-3), 5.02 (dd, J = 10.3, 3.9 Hz, 1 H, H-4), 4.40 (q, J = 6.5 Hz, 1 H, H-5), 2.17, 2.10, 2.00 (3 x s, 3 x 3 H, 3 x OAc), 1.21 (d, J = 6.5 Hz, 3 H, H-6); **¹³C NMR** (CDCl₃, 50 MHz) δ 170.93, 170.72, 170.68 (3 x CO), 84.08 (C-1), 73.47 (C-5), 72.49 (C-3), 70.11 (C-4), 67.66 (C-2), 20.09, 20.08, 20.04 (3 x Ac), 14.86 (C-6); **MS** (ESI⁻) m/z = 352.15 [M-H]; $[\alpha]_D^{20}$ = +9.13 (c = 1.0, CDCl₃); **IR** (ATR) 1766.91 (CO).

1-ethylthio-2,3,4-tri-O-acetyl-6-deoxy- β -L-galactopyranoside **2.18**^[16]



To a solution of **2.17** (31.25 g, 88.5 mmol, 1 eq.) in ACN (180 mL), was added thiourea (7.07 g, 92.9 mmol, 1.05 eq.) and the mixture was refluxed at 80 °C for 30 min. The reaction was filtered, the precipitate washed with ACN and re-suspended in ACN (180 mL). Triethylamine (30.8 mL, 221.25 mmol, 2.5 eq.) and EtBr (9.9 mL, 132.75 mmol, 1.5 eq.) were added, and the reaction stirred at r.t for 1 h. The mixture was filtered, the precipitate dissolved in DCM (150 mL) and washed with H₂O (125 mL), HCl 1M (125 mL) and brine (125 mL), dried over MgSO₄ and evaporated to afford the title compound **2.18** as a colourless oil (22.20 g, 66.4 mmol, 75%). **TLC** R_f = 0.9 (EtOAc:PE, 4:1); **¹H NMR** (CDCl₃, 200 MHz) δ 5.25 (ddd, J = 18.7, 11.0, 6.5 Hz, 2 H, H-4, H-2), 5.04 (dd, J = 9.9, 3.4 Hz, 1 H, H-3), 4.45 (d, J = 9.7 Hz, 1 H, H-1), 3.97-3.61 (m, 1 H, H-5), 2.73 (qd, J = 7.4, 3.4 Hz, 2 H, CH₂), 2.17, 2.06, 1.98 (3 x s, 3 x 3 H, 3 x OAc), 1.29 (d, J = 7.4 Hz, 3 H, H-6), 1.21 (d, J = 6.4 Hz, 3 H, CH₃); **¹³C NMR** (CDCl₃, 50 MHz) δ 170.63, 170.52, 170.33 (3 x CO), 83.66 (C-1), 73.33 (C-5), 72.50 (C-3), 70.61 (C-4), 67.52 (C-2), 24.25 (CH₂), 20.03, 20.00, 19.98 (3 x Ac), 16.57 (CH₃), 14.87 (C-6); **MS** (ESI⁻) m/z = 333.37 [M-H]; $[\alpha]_D^{20}$ = +7.40 (c = 1.0, CDCl₃);^[17] **IR** (ATR) 1745.70 (CO).

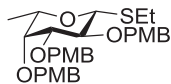
1-ethylthio-6-deoxy- β -L-galactopyranoside **2.19**



To a solution of **2.18** (22.2 g, 66.4 mmol, 1 eq.) in MeOH (66 mL), was added NaOMe (25% in MeOH, 13.6 mL, 66.4 mmol, 1 eq.) and the mixture was stirred at r.t. for 1 h. The reaction was quenched with Dowex H⁺, filtered and evaporated to afford the title compound **2.19** as a white solid (15.80 g, mmol, quant.). **TLC** R_f = 0.0 (EtOAc:PE, 4:1); **¹H NMR** (MeOD, 200 MHz) δ 4.30 (dd, J = 8.5, 0.9 Hz, 1 H, H-1), 3.73-3.39 (m, 4 H, H-2, H-3, H-4, H-5), 2.86-2.57 (m, 2 H, CH₂), 1.25 (d, J = 6.5 Hz, 3 H, CH₃); **¹³C NMR** (MeOD, 50 MHz) δ 84.68 (C-1), 71.92 (C-5), 73.04 (C-3), 70.58 (C-4), 66.54 (C-2), 23.99 (CH₂), 16.52 (CH₃), 14.47 (C-6); **MS** (ESI⁻) m/z = 207.27 [M-H]; $[\alpha]_D^{20}$ = +33.72 (c = 1.0, CDCl₃); **IR** (ATR) 3401.76 (OH).

1-ethylthio-2,3,4-tri-O-paramethoxybenzyl-6-deoxy- β -L-galactopyranoside

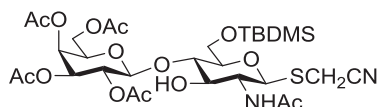
2.20



To a solution of **2.19** (15.80 g, 75.85 mmol, 1 eq.) and TBAI (2.8 g, 7.58 mmol, 0.1 eq.) in DMF (dry, 150 mL), was added, at 0 °C and portion wise NaH (11 g, 455.1 mmol, 6 eq.). The mixture was stirred at 0 °C for 1 h, then PMBCl (41.1 mL, 303.4 mmol, 4 eq.) was added and the reaction stirred at r.t. for 16 h. The reaction was carefully quenched at 0 °C with EtOH, then evaporated, recovered with EtOAc (200 mL) and washed with brine (4 x 250 mL). The organic layer was dried over MgSO₄, evaporated and re-crystallised from hot IPA to afford the title compound **2.20** as a white solid (21 g, 36.9 mmol, 49%). **TLC** R_f = 0.8 (EtOAc:PE, 4:1); **¹H NMR** (CDCl₃, 400 MHz) δ 7.45-6.66 (m, 12 H, PMB), 4.92-4.64 (m, 6 H, 3 x OCH₂), 4.38 (d, J = 9.6 Hz, 1 H, H-1), 3.84 1 3.80 (m, 9 H, 3 x OMe), 3.77 (d, J = 9.5 Hz, 1 H, H-2), 3.57 (d, J = 2.6 Hz, 1 H, H-4), 3.53 (dd, J = 9.2, 2.8 Hz, 1 H, H-3), 3.45 (q, J = 6.3 Hz, 1 H, H-5), 2.85-2.61 (m, 2 H, SCH₂), 1.30 (t, J = 7.4 Hz, 3 H, SCH₂CH₃), 1.17 (d, J = 6.3 Hz, 3 H, H-6); **¹³C NMR** (CDCl₃, 50 MHz) δ 159.97, 159.46, 159.23 (3 x COMe-Ar), 130.75, 130.64, 130.22 (3 x CCH₂-Ar), 129.94, 129.47, 129.12, 113.42, 113.40, 113.28 (12 x CH-Ar), 84.68 (C-1), 74.48, 74.33, 74.29 (3 x CH₂), 71.92 (C-5), 73.04 (C-3), 70.58 (C-4), 66.54 (C-2), 55.09, 54.98, 54.67 (3 x OMe), 23.99 (CH₂), 16.52 (CH₃), 14.47 (C-6); **MS** (ESI⁻) m/z = 567.71 [M-H]; $[\alpha]_D^{20}$ = -2.13 (c = 1.0, CDCl₃); **IR** (ATR) 819.66 (arom).

2-N-acetamido-6-O-*tert*-butyl-dimethylsilyl-4-O-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside

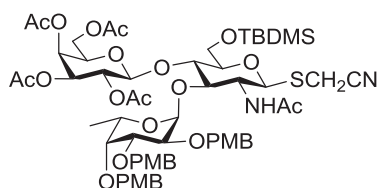
2.21



A mixture of **2.5** (1 g, 2.6 mmol, 1.0 eq.) and **2.12** (1.4 g, 2.8 mmol, 1.1 eq.) was co-evaporated with toluene, dried under high vacuum and suspended in anhydrous DCM (7 mL) with MS (4 Å, 2.5 g) and stirred for 1 h. The mixture was cooled to -41 °C and stirred for 15 min prior to addition of boron trifluoride diethyl etherate (370 μ L, 3.0 mmol, 1.15 eq.). The solution was stirred at -41 °C for 1 h before triethylamine (1 mL) was added and the crude mixture was directly purified by flash column chromatography over silica (EtOAc:PE, 1:1 to 1:0) to afford the title compound **2.21** as a white solid (776.7 mg, 1.076 mmol, 42%). **TLC** R_f = 0.5 (EtOAc, 100%); **$^1\text{H NMR}$** (CDCl_3 , 400 MHz) δ 5.97 (d, J = 8.9 Hz, 1 H, NH), 5.42-5.32 (m, 1 H, H-4'), 5.20 (dd, J = 10.5, 8.0 Hz, 1 H, H-2'), 4.96 (dd, J = 10.5, 3.4 Hz, 1 H, H-3'), 4.67 (d, J = 10.3 Hz, 1 H, H-1), 4.61 (d, J = 8.1 Hz, 1 H, H-1'), 4.24-4.05 (m, 2 H, H-6'), 3.99 (dd, J = 7.4, 6.3 Hz, 1 H, H-5'), 3.86 (dd, J = 20.0, 10.1 Hz, 2 H, H-2, H-6a), 3.70 (ddd, J = 21.0, 10.0, 5.1 Hz, 3 H, H-6b, H-3, H-4), 3.56 (d, J = 16.9 Hz, 1 H, CH_2aCN), 3.38 (dd, J = 9.1, 1.8 Hz, 1 H, H-5), 3.32 (d, J = 16.9 Hz, 1 H, CH_2bCN), 2.14, 2.06, 2.04, 2.01, 1.96 (5 x s, 5 x 3 H, 5 x Ac), 0.90 (s, 9 H, ^tBu), 0.07, 0.06 (2 x s, 2 x 3 H, 2 x CH_3); **$^{13}\text{C NMR}$** (CDCl_3 , 101 MHz) δ 170.90, 170.53, 170.09, 170.00, 169.26 (5 x CO), 116.38 (CN), 101.46 (C-1'), 82.61 (C-1), 80.32 (C-5), 78.84 (C-3), 73.41 (C-5'), 71.32 (C-3'), 70.80 (C-4'), 68.65 (C-4), 66.77 (C-2'), 61.32 (C-6'), 60.42 (C-6), 54.19 (C-2), 25.82 (C(CH_3)₃),

23.32 (NHAc), 20.70, 20.60, 20.55, 20.52 (4 x OAc), 18.21 ($\text{C}(\text{CH}_3)_3$), 14.31 (CH_2CN), -4.98, -5.22 (2 x CH_3); **MS** (ESI⁺) m/z = 721.87 [M+H]⁺; $[\alpha]_{\text{D}}^{20}$ = -43.27 (c = 1.0, CHCl_3); **IR** (ATR) 3347.95 (OH), 2201.76 (CN), 1732.17 (CO), 1693.87 (CONH).

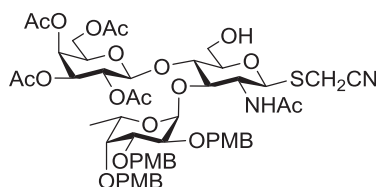
2-N-acetamido-6-O-*tert*-butyl-dimethylsilyl-3-O-[2,3,4-tri-O-paramethoxybenzyl-6-deoxy- α -L-galactopyranosyl]-4-O-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside 2.22



2.21 (380 mg, 0.527 mmol, 1.0 eq.) and **2.20** (599.6 mg, 1.054 mmol, 2 eq.), co-evaporated with toluene and dried under high vacuum for 1h, were dissolved in DCM:DMF (dry, 1:1, 1.5 mL). To the mixture were added MS (4 Å, 790 mg) and it was stirred for 1 h at r.t. before copper(II) bromide (235.4 mg, 1.054 mmol, 2.0 eq.) and tetra-butylammonium bromide (356.8 mg, 1.107 mmol, 2.1 eq.) were added too. The mixture was stirred at r.t. for 16 h in the dark, then filtered over Celite and washed with DCM (200 mL). The filtrate was washed with brine (150 mL) and NaHCO_3 satd. solution (6 x 120 mL). The aqueous layer was re-extracted with DCM (200 mL) and the combined organic layers were dried over MgSO_4 , filtered, concentrated in vacuum and directly purified by flash column

chromatography over silica (PE:EtOAc, 1:1 to 7:3) to afford the title compound **2.22** as a white foam (513.3 mg, 0.418 mmol, 79%). **TLC** R_f = 0.1 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 7.35-6.77 (m, 12 H, CH-arom), 6.22 (d, J = 7.9 Hz, 1 H, NH), 5.36 (dd, J = 3.4, 0.8 Hz, 1 H, H-4'), 5.07 (dd, J = 10.6, 7.9 Hz, 2 H, H-2', H-1''), 4.93 (dd, J = 10.5, 3.5 Hz, 1 H, H-3'), 4.88 (d, J = 11.4 Hz, 1 H, CH₂-PMB), 4.77 (dd, J = 9.4, 8.0 Hz, 3 H, CH₂-PMB, H-1, H-1'), 4.70 (s, 1 H, CH₂-PMB), 4.66 (d, J = 11.1 Hz, 1 H, CH₂-PMB), 4.61 (d, J = 11.5 Hz, 1 H, CH₂-PMB), 4.33 (d, J = 6.4 Hz, 1 H, H-5''), 4.17 (dd, J = 11.0, 7.8 Hz, 1 H, H-6'a), 4.11-4.05 (m, 1 H, H-2''), 4.04-3.97 (m, 3 H, H-6'b, H-3, H-5), 3.89-3.76 (m, 14 H, H-2, H-6, H-5', H-3'', 3 x OCH₃), 3.65 (d, J = 1.6 Hz, 1 H, H-4''), 3.45 (d, J = 17.1 Hz, 1 H, CH₂aCN), 3.40-3.32 (m, 1 H, H-4), 3.14 (d, J = 17.0 Hz, 1 H, CH₂bCN), 2.05, 2.01, 1.96, 1.94 (4 x s, 4 x 3 H, 4 x OAc), 1.77 (s, 3 H, NHAc), 1.15 (d, J = 6.5 Hz, 3 H, H-6''), 0.90 (s, 9 H, ^tBu), 0.07, 0.05 (2 x s, 2 x 3 H, 2 x CH₃); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.50, 170.12, 170.03, 170.00, 169.42 (5 x CO), 159.36, 159.17, 159.02 (3 x COCH₃-arom), 130.81, 130.65, 130.52 (3 x C-arom), 129.99, 129.95, 128.77 (3 x CHC-arom), 116.80 (CN), 113.93, 113.76, 113.58 (3 x CHCO-arom), 99.19 (C-1'), 98.34 (C-1''), 82.02 (C-1), 79.97 (C-4), 79.61 (C-3''), 76.48, 76.27, 76.04 (C-4'', C-2'', C-3), 74.01, 73.78 (2 x OCH₂), 72.97 (C-5), 72.21 (OCH₂), 71.29 (C-3'), 70.69 (C-5'), 68.97 (C-2'), 66.79 (C-5''), 61.18 (C-4'), 60.57 (C-6), 60.39 (C-6'), 55.28, 55.25, 55.22 (3 x OCH₃), 55.18 (C-2), 25.84 (C(CH₃)₃), 23.09 (NHAc), 21.05, 20.83, 20.61, 20.55 (4 x OAc), 18.17 (C(CH₃)₃), 16.78 (C-6''), 14.22 (CH₂CN), -5.07, -5.33 (2 x SiCH₃); **MS** (ESI⁺) m/z = 1228.45 [M+H]; $[\alpha]_D^{20}$ = -60.34 (c = 1.0, CHCl₃); **IR** (ATR) 2211.43 (CN), 1716.94 (CO), 1696.88 (CONH).

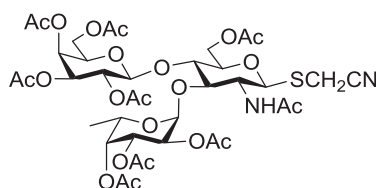
2-N-acetamido-3-O-[2,3,4-tri-O-paramethoxybenzyl-6-deoxy- α -L-galactopyranosyl]-4-O-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.23**



To **2.22** (483.5 mg, 0.394 mmol, 1.0 eq) dissolved in THF (dry, 3.5 mL) was added triethylamine trihydrofluoride (460 μ L, 2.82 mmol, 7 eq). The solution was stirred at r.t. for 16 h. The solvent was evaporated and the crude mixture was purified by flash column chromatography on silica (EtOAc:PE, 7:3 to 9:1) to afford the title compound **2.23** as a white powder (366.7 mg, 0.329 mmol, 84%). **TLC** R_f = 0.2 (PE:EtOAc, 1:9); **$^1\text{H-NMR}$** (CDCl_3 , 400 MHz) δ 7.32-7.11 (m, 6 H, 6 x CH-PMB), 6.91-6.65 (m, 6 H, 6 x CH-PMB), 6.36 (d, J = 7.7 Hz, 1 H, NH), 5.29 (d, J = 2.7 Hz, 1 H, H-4'), 5.08 (d, J = 3.5 Hz, 1 H, H-1''), 4.98 (ddd, J = 13.9, 10.5, 5.7 Hz, 2 H, H-2', H-3'), 4.80-4.50 (m, 8 H, H-1', H-1, 3 x CH_2), 4.29 (d, J = 6.5 Hz, 1 H, H-5''), 4.16-3.55 (m, 21 H, H-4'', H-6', H-2'', H-3, H-4, H-2, H-6, H-5', H-3', CH_2aCN , 3 x OMe), 3.33 (d, J = 7.6 Hz, 1 H, H-5), 3.06 (d, J = 17.4 Hz, 1 H, CH_2aCN), 1.99, 1.92, 1.91, 1.90 (4 x s, 4 x 3 H, 4 x OAc), 1.71 (s, 3 H, NHAc), 1.12 (d, J = 6.5 Hz, 3 H, H-6''); **$^{13}\text{C-NMR}$** (CDCl_3 , 63 MHz) δ 171.69, 171.38, 170.64, 170.44, 169.76 (5 x CO), 159.78, 159.61, 159.44 (3 x CO-arom), 131.19, 131.04, 130.78 (3 x CC-arom), 130.41, 130.38, 129.11 (3 x CHCC-arom), 118.42 (CN), 114.32, 114.18, 114.02 (3 x CHCO-arom), 100.34 (C-1'), 98.57 (C-1''), 83.70 (C-1), 80.34 (C-5), 80.07 (C-3''), 76.83, 76.72, 76.48 (C-4'', C-2'', C-3),

74.41, 74.07, 74.02, 72.43 (C-4, 3 x CH₂), 71.17 (C-5'), 71.06 (C-3'), 69.64 (C-2'), 67.25 (C-5''), 67.04 (C-4'), 61.06 (C-6), 60.85 (C-6'), 55.70, 55.68, 55.63 (3 x OMe), 55.06 (C-2), 23.55 (NHAc), 21.48, 21.22, 21.01, 20.97 (4 x OAc), 17.28 (C-6''), 14.60 (CH₂CN); **MS** (ESI⁺) *m/z* = 1135.3 [M+Na]; [α]_D²⁰ = -60.3 (c = 1.1, CHCl₃); **IR** (ATR) 3401.00 (OH), 2211.63 (CN), 1719.82 (CO), 1696.28 (CONH).

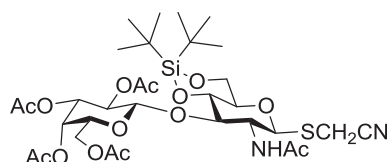
2-N-acetamido-6-O-acetyl-3-O-[2,3,4-tri-O-acetyl-6-deoxy- α -L-galactopyranosyl]-4-O-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.24**



To **2.23** (647 mg, 0.58 mmol, 1.0 eq.) dissolved in ACN:H₂O (9:1, 10 mL) was added CAN (2.87 g, 5.23 mmol, 9 eq.). The solution was stirred at r.t. for 2 h, then Ac₂O (30 mL), pyridine (60 mL) and DMAP (7.1 mg, 0.058 mmol, 0.1 eq.) were added. The reaction was stirred at r.t. for 16 h. The solvents were evaporated, the crude mixture was co-evaporated with toluene and purified by flash column chromatography on silica (EtOAc:PE, 1:1 to 1:0) to afford the title compound **2.24** as a white powder (325 mg, 0.34 mmol, 59%). **TLC** *R_f* = 0.5 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 5.86 (d, *J* = 9.8 Hz, 1 H, NH), 5.44 (d, *J* = 4.0 Hz, 1 H, H-1''), 5.42 (d, *J* = 3.4 Hz, 1 H, H-4'), 5.39 (d, *J* = 2.7 Hz, 1 H, H-4''), 5.18 (dd, *J* = 11.0, 3.3 Hz, 1 H, H-3''), 5.09 (dd, *J* = 10.3, 8.0 Hz, 1 H, H-2'), 5.00 (dd, *J* =

10.3, 3.7 Hz, 1 H, H-3'), 5.00 (dd, $J = 10.0, 3.5$ Hz, 1 H, H-2''), 4.92 (q, $J = 6.5$ Hz, 1 H, H-5''), 4.70 (dd, $J = 12.2, 2.1$ Hz, 2 H, H-1, H-6a), 4.52 (dd, $J = 16.5, 7.1$ Hz, 2 H, H-6'a, H-1'), 4.32 (dd, $J = 11.5, 7.9$ Hz, 1 H, H-6'b), 4.07 (m, 2 H, H-6b, H-2), 3.94-3.85 (m, 3 H, H-, H-3, H-5'), 3.61 (d, $J = 17.1$ Hz, 1 H, CH₂aCN), 3.55 (m, 1 H, H-5), 3.28 (d, $J = 17.1$ Hz, 1 H, CH₂bCN), 2.20, 2.15, 2.15, 2.10, 2.09, 2.07, 1.99, 1.98, 1.97 (9 x s, 9 x 3 H, 9 x Ac), 1.21 (d, 3 H, $J = 5.0$, H-6''); ¹³C NMR (101 MHz, CDCl₃) δ 171.42, 170.90, 170.78, 170.66, 170.50, 170.31, 170.08, 169.84, 169.16 (9 x CO), 116.42 (CN), 100.59 (C-1'), 95.64 (C-1''), 82.97 (C-1), 77.58 (C-5), 74.90 (C-3), 74.24 (C-4), 71.34 (C-4''), 71.20 (C-2''), 70.97 (C-2), 68.95 (C-2'), 68.06 (C-3'), 67.91 (C-3''), 66.67 (C-4'), 64.49 (C-5'), 64.43 (C-5''), 61.41 (C-6), 60.69 (C-6'), 25.40 (C-6''), 23.38 (NHAc), 20.97, 20.93, 20.90, 20.86, 20.78, 20.70, 20.65, 20.57 (8 x Ac), 14.32 (CH₂CN); MS (ESI⁺) m/z = 921.28 [M+H]; $[\alpha]_D^{20} = -123.8$ ($c = 1.0$, CHCl₃); IR (ATR) 2206.78 (CN), 1743.13 (CO).

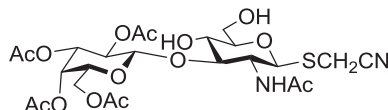
2-N-acetamido-4,6-O-di-*tert*-butyl-dimethylsilylidene-3-O-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside 2.25



2.6 (1.0 g, 2.4 mmol, 1 eq.) and **2.10** (1.48 g, 3.6 mmol, 1.5 eq.) were co-evaporated with toluene and dried under high vacuum for 1 h. DCM (dry, 24 mL)

was added, the mixture transferred into a flame-dried flask with MS (4Å, 2.4 g) and stirred at r.t. for 1 h before adding DTBMP (493 mg, 2.4 mmol, 1 eq.) and AgOTf (1.418 g, 5.5 mmol, 2.3 eq.). The reaction was stirred at r.t. for 16 h in the dark. The reaction mixture was filter over Celite and rinsed with DCM. After removal of the solvent, the residue was purified by flash column chromatography on silica (EtOAc:PE, 1:1 to 1:0) to afford the title compound **2.25** as a white amorphous solid (1.453 g, 2.01 mmol, 84%). **TLC** R_f = 0.25 (PE:EtOAc, 1:2); **¹H NMR** (CDCl₃, 400 MHz) δ 5.73 (d, J = 8.2 Hz, 1 H, NH), 5.36 (dd, J = 3.3, 0.8 Hz, 1 H, H-4'), 5.16 (dd, J = 10.4, 7.7 Hz, 1 H, H-2'), 5.06 (d, J = 10.4 Hz, 1 H, H-1), 5.00-4.94 (m, 2 H, H-3', H-1'), 4.20 (dd, J = 10.3, 5.0 Hz, 1 H, H-6a), 4.15-4.08 (m, 3 H, H-6', H-3), 4.05 (d, J = 9.4 Hz, 1 H, H-6b), 3.96-3.83 (m, 3 H, H-4, H-5', H-2), 3.56-3.49 (m, 1 H, H-5), 3.51 (d, J = 17.2 Hz, 1 H, CH₂aCN), 3.33 (d, J = 17.2 Hz, 1 H, CH₂bCN), 2.14, 2.09, 2.04, 2.03, 1.98 (5 x Ac), 1.07, 1.00 (2 x s, 2 x 9 H, 2 x ^tBu); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.76, 170.37, 170.25, 170.24, 169.85 (5 x CO), 116.74 (CN), 99.76 (C-1'), 82.45 (C-1), 79.59 (C-3), 75.81 (C-4), 75.04 (C-5), 71.15 (C-3'), 70.74 (C-5'), 70.14 (C-2'), 66.76 (C-4'), 66.10 (C-6), 61.29 (C-6'), 55.21 (C-2), 27.32, 26.98 (2 x C(CH₃)₃), 23.51 (2 x C(CH₃)₃), 22.67 (NHAc), 20.91, 20.77, 20.61, 19.96 (4 x OAc), 14.30 (CH₂CN); **MS** (ESI⁺) m/z = 721.3 [M+H]; $[\alpha]_D^{20}$ = -21.4 (c = 1, CHCl₃); **IR** (ATR) 2211.34 (CN), 1718.63 (CO), 1698.01 (CONH).

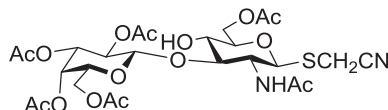
2-N-acetamido-3-O-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.26**



To **2.25** (1.445 g, 1.93 mmol, 1 eq.) in THF (dry, 20 mL) was added triethylamine hydrofluoride (662 μ L, 4.06 mmol, 2.1 eq.). After 1 h the solvent was evaporated and the residue co-evaporated with MeOH and toluene, to afford the title compound **2.26** as a white amorphous solid (1.717 g, 1.9 mmol, quant.). **TLC** R_f = 0.0 (PE:EtOAc, 1:2); **^1H NMR** (CD_3CN , 400 MHz) δ 6.88 (d, J = 9.5 Hz, 1 H, NH), 5.34 (d, J = 2.3 Hz, 1 H, H-4'), 5.05 (dd, J = 5.2, 3.5 Hz, 2 H, H-2', H-3'), 4.81-4.74 (d, J = 10.4 Hz, 1 H, H-1'), 4.68 (d, J = 10.4 Hz, 1 H, H-1), 4.17-4.12 (m, 2 H, H-5', H-6a'), 4.11-4.05 (m, 1 H, H-6b'), 3.79 (m, 2 H, H-2, H-6a), 3.73-3.67 (m, 2 H, H-6b, H-3), 3.67 (d, J = 11.7 Hz, 1 H, CH_2aCN), 3.61 (d, J = 12.1 Hz, 1 H, CH_2bCN), 3.39-3.33 (m, 2 H, H-4, H-5), 2.10, 2.01, 1.91, 1.91 (5 x s, 5 x 3 H, 5 x Ac); **^{13}C NMR** (CD_3CN , 63 MHz) δ 171.19, 171.01, 170.75, 170.51 (5 x CO), 118.81 (CN), 101.84 (C-1'), 85.46 (C-3), 84.18 (C-1), 81.36 (C-5), 71.96 (C-5'), 71.58 (C-3'), 69.86 (C-4), 69.42 (C-2'), 68.17 (C-4'), 62.71 (C-6), 62.35 (C-6'), 53.56 (C-2), 23.33, 20.94, 20.81, 20.73, 20.69 (5 x Ac), 15.37 (CH_2CN); **MS** (ESI $^-$) m/z = 605.59 [M-H]; $[\alpha]_D^{20}$ = -42.5 (c = 1, acetone); **IR** (ATR) 3398.72 (OH), 2210.32 (CN), 1716.23 (CO), 1699.11 (CONH).

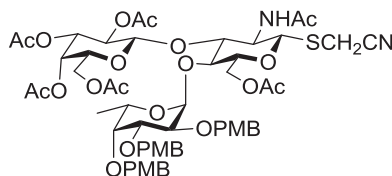
2-N-acetamido-6-O-acetyl-3-O-[2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy-β-D-glucopyranoside

2.27



To **2.26** (1.71 g, 1.93 mmol, 1 eq.) in DCM (dry, 20 mL) and pyridine (dry, 624 μ L, 7.72 mmol, 4 eq.) under argon, at -78°C , was added AcCl (343 μ L, 4.825 mmol, 2.5 eq.). After 1 h MeOH was added to quench the reaction and stirred for a few min at -78°C . The crude was evaporated, recovered with EtOAc (100 mL) and washed with HCl 1M (100 mL), NaHCO_3 (100 mL), brine (100 mL) and dried over MgSO_4 . Evaporation of the solvent afforded the title compound **2.27** as a white amorphous solid (294 mg, 0.45 mmol, 23%). **TLC** R_f = 0.5 (MeOH:EtOAc, 0.5:10); **$^1\text{H NMR}$** (CDCl_3 , 400 MHz) δ 5.57 (d, J = 8.1 Hz, 1 H, NH), 5.36 (d, J = 3.5 Hz, 1 H, H-4'), 5.20 (dd, J = 10.5, 8.0 Hz, 1 H, H-2'), 5.03-4.93 (m, 2 H, H-3', H-1), 4.54 (d, J = 8.0 Hz, 1 H, H-1'), 4.11-3.94 (m, 3 H, H-6, H-6'a), 3.92-3.89 (m, 2 H, H-6'b, H-5'), 3.76 (dd, J = 12.1, 5.3 Hz, 1 H, H-3), 3.63-3.44 (m, 4 H, H-2, H-4, H-5, CH_2aCN), 3.34 (d, J = 17.3 Hz, 1 H, CH_2bCN), 2.14, 2.05, 2.04, 2.02, 2.01, 1.96 (6 x s, 6 x 3 H, 6 x Ac); **$^{13}\text{C NMR}$** (CDCl_3 , 101 MHz) δ 171.06, 171.00, 170.77, 170.63, 170.21, 169.43 (6 x CO), 116.91 (CN), 101.58 (C-1'), 84.58 (C-3), 81.75 (C-1), 78.22 (C-5), 71.29 (C-5'), 70.78 (C-3'), 69.16 (C-4), 68.89 (C-2'), 67.00 (C-4'), 63.24 (C-6), 61.53 (C-6'), 54.39 (C-2), 23.62, 21.09, 21.03, 20.84, 20.72, 20.64 (6 x Ac), 14.32 (CH_2CN); **MS** (ESI^-) m/z = 647.63 [M-H]; $[\alpha]_D^{20}$ = -15.4 (c = 1, CHCl_3); **IR** (ATR) 3396.44 (OH), 2212.75 (CN), 1716.94 (CO), 1695.38 (CONH).

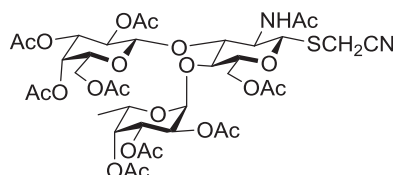
2-N-acetamido-6-O-acetyl-4-O-[2,3,4-tri-O-paramethoxybenzyl-6-deoxy- α -L-galactopyranosyl]-3-O-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.28**



2.27 (294 mg, 450 μ mol, 1.0 eq.) and **2.20** (512 mg, 900 μ mol, 2.0 eq.) co-evaporated with toluene and dried under high vacuum for 1 h, were dissolved in DCM:DMF (dry, 1:1, 2 mL). MS (4Å, 450 mg) were added and the reaction was stirred 1 h at r.t. before copper(II) bromide (201 mg, 900 μ mol, 2.0 eq.) and tetrabutylammonium bromide (306 mg, 950 μ mol, 2.1 eq.) were added. The mixture was stirred at r.t. for 16 h with light exclusion. The crude mixture was filtered over Celite and washed with EtOAc (200 mL). The filtrate was washed with brine (200 mL) and NH_4Cl satd. solution (pH 9, 200 mL). The organic layer was dried over MgSO_4 , filtered, concentrated under vacuum and purified by flash column chromatography on silica (EtOAc:PE, 1:2 to 1:0) to afford the title compound **2.28** as a white foam (520.2 mg, 0.45 mmol, quant.). **TLC** R_f = 0.15 (PE:EtOAc, 2:1); **^1H NMR** (CDCl_3 , 400 MHz) δ 7.39-7.09 (m, 6 H, CH-PMB), 6.97-6.70 (m, 6 H, CH-PMB), 5.34 (d, J = 3.4 Hz, 1 H, NH), 5.17-5.07 (m, 2 H, H-4', H-2'), 4.96-4.86 (m, 3 H, H-3', H-1'', CH_2 -PMB), 4.80-4.65 (m, 7 H, H-1, H-1', CH_2 -PMB), 4.60-4.50 (m, 3 H, H-6, H-5''), 4.09-4.01 (m, 4 H, H-2'', H-2, H-6'), 3.99-3.76 (m, 16 H, H-3, H-4, H-5', 3 x OMe, H-5, H-4'', CH_2CN), 2.07, 2.04, 2.00, 1.96, 1.92, 1.86 (6 x s, 6 x 3 H, 6 x Ac); **^{13}C NMR** (CDCl_3 , 101 MHz) δ 170.62, 170.39, 170.29,

170.05, 169.97, 169.63 (6 x CO), 159.61, 159.41, 159.27 (3 x CCO-arom), 130.84, 130.66 (4 x CC-arom), 130.56, 130.23 (2 x CHCC-arom), 130.07 (2 x CC-arom), 128.77 (CHCC-arom), 116.92 (CN), 114.04, 113.97, 113.79 (6 x CHCO-arom), 100.79 (C-1'), 96.91 (C-1''), 81.72 (C-1), 80.26 (C-3''), 77.92 (C-3), 76.73 (C-5), 76.03 (C-4''), 75.09 (C-2''), 74.54, 73.89, 72.33 (3 x CH₂), 71.49 (C-4), 71.06 (C-3'), 70.57 (C-5'), 68.28 (C-2'), 67.00 (C-5''), 66.79 (C-4'), 62.24 (C-6), 60.34 (C-6'), 55.41, 55.40, 55.38 (3 x OMe), 53.52 (C-2), 23.33 (NHAc), 21.05, 20.88, 20.80, 20.67, 20.61 (5 x OAc), 16.93 (C-6''), 14.57 (CH₂CN); **MS** (ESI⁺) *m/z* = 1156.23 [M+H]⁺; [α]_D²⁰ = -51.7 (c = 1, CHCl₃); **IR** (ATR) 2212.48 (CN), 1718.25 (CO), 1694.99 (CONH).

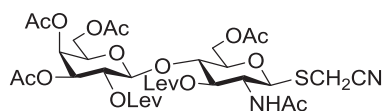
2-N-acetamido-6-O-acetyl-4-O-[2,3,4-tri-O-acetyl-6-deoxy-α-L-galactopyranosyl]-3-O-[2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy-β-D-glucopyranoside 2.29



To **2.28** (521.2 mg, 0.45 mmol, 1.0 eq.) dissolved in ACN:H₂O (9:1, 10 mL) was added CAN (1.5 g, 2.72 mmol, 6.05 eq.). The solution was stirred at r.t. for 2 h, then Ac₂O (45 mL), pyridine (45 mL) and DMAP (5.5 mg, 0.045 mmol, 0.1 eq.) were added. The reaction was stirred at r.t. for 16 h. Solvents were evaporated, the crude mixture was co-evaporated with toluene and purified by flash column

chromatography on silica (EtOAc:PE, 1:1 to 1:0) to afford the title compound **2.29** as a white powder (222.6 mg, 0.24 mmol, 54%). **TLC** R_f = 0.2 (PE:EtOAc, 1:9); **¹H NMR** (CDCl₃, 400 MHz) δ 5.86 (d, J = 9.6 Hz, 1 H, NH), 5.42 (d, J = 2.8 Hz, 1 H, H-4'), 5.39-5.35 (m, 1 H, H-4''), 5.22 (qd, J = 11.0, 3.3 Hz, 2 H, H-2'', H-3''), 5.12 (dd, J = 10.4, 8.2 Hz, 1 H, H-2'), 4.97 (ddd, J = 13.9, 12.0, 3.5 Hz, 3 H, H-1'', H-5'', H-3'), 4.76 (d, J = 8.2 Hz, 1 H, H-1'), 4.71 (d, J = 10.2 Hz, 1 H, H-1), 4.58 (dd, J = 12.3, 2.2 Hz, 1 H, H-6a), 4.50 (dd, J = 11.4, 5.8 Hz, 1 H, H-6'a), 4.25 (dd, J = 11.4, 8.4 Hz, 1 H, H-6'b), 4.14 (dd, J = 19.2, 10.0 Hz, 1 H, H-2), 3.99 (t, J = 9.1 Hz, 1 H, H-3), 3.95-3.79 (m, 3 H, H-5', H-6b, H-4), 3.73-3.55 (m, 2 H, CH₂aCN, H-5), 3.31 (d, J = 17.1 Hz, 1 H, CH₂bCN), 2.19, 2.17, 2.10, 2.09, 2.09, 2.07, 2.06, 1.98, 1.96 (9 x s, 9 x 3 H, 9 x Ac), 1.26 (d, J = 6.6 Hz, 3 H, H-6''); **¹³C NMR** (CDCl₃, 126 MHz) δ 171.03, 170.77, 170.69, 170.66, 170.63, 170.59, 170.53, 170.40, 170.36 (9 x CO), 113.74 (CN), 101.27 (C-1'), 97.23 (C-1''), 82.34 (C-1), 78.86 (C-3), 78.45 (C-5), 71.52 (C-4), 71.28 (C-4''), 71.05 (C-3'), 70.48 (C-5'), 68.50 (C-2'), 68.36 (C-3''), 67.89 (C-2''), 66.70 (C-4'), 64.57 (C-5''), 60.54 (C-6), 60.16 (C-6'), 55.30 (C-2), 23.45 (NHAc), 20.94, 20.79, 20.75, 20.72, 20.69, 20.62, 20.56, 20.42 (8 x OAc), 15.86 (C-6''), 14.15 (CH₂CN); **MS** (ESI⁺) m/z = 943.26 [M+Na]; $[\alpha]_D^{20}$ = -118.8 (c = 1, CHCl₃); **IR** (ATR) 2208.38 (CN), 1747.19 (CO).

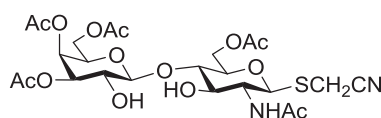
2-N-acetamido-6-O-acetyl-4-O-[3,4,6-tri-O-acetyl-2-O-levulinoyl- β -D-galactopyranosyl]-3-O-levulinoyl-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.30**



2.9 (1.0 g, 2.4 mmol, 1 eq.) and **2.15** (1.682 g, 3.6 mmol, 1.5 eq) were co-evaporated with toluene and dried under high vacuum for 1 h. DCM:toluene (dry, 1:1, 24 mL) were added and the mixture was transferred to a flame-dried flask and stirred at -45 °C for 30 min before addition of AgOTf (1.048 g, 4.08 mmol, 1.7 eq.). The temperature was raised up to 0 °C in 4 h and the reaction quenched with triethylamine (3 mL), filtered over Celite and rinsed with DCM. After removal of the solvent, the residue was recovered with DCM (100 mL), washed with H₂O (100 mL), KI satd. solution (100 mL), brine (100 mL), dried over MgSO₄, evaporated and purified by flash column chromatography on silica (EtOH:CHCl₃, 0.25:10) to afford the title compound **2.30** as a white foam (710 mg, 0.88 mmol, 37%). **TLC** R_f = 0.15 (EtOH:CHCl₃, 0.5:10); **¹H NMR** (CDCl₃, 400 MHz) δ 5.34 (d, *J* = 3.0 Hz, 1 H, NH), 5.17-5.03 (m, 2 H, H-4', H-3), 5.00-4.95 (m, 1 H, H-2'), 4.76 (dd, *J* = 10.2, 6.0 Hz, 1 H, H-3'), 4.56-4.50 (m, 1 H, H-1), 4.48 (d, *J* = 7.8 Hz, 1 H, H-6a), 4.16-4.02 (m, 5 H, H-6b, H-1', H-2, H-6'a), 3.89 (t, *J* = 6.8 Hz, 1 H, H-5'), 3.84-3.74 (m, 2 H, H-4, H-5), 3.65 (dd, *J* = 17.1, 1.9 Hz, 1 H, CH₂aCN), 3.28 (dd, *J* = 17.0, 4.4 Hz, 1 H, CH₂bCN), 2.83-2.64 (m, 4 H, 2 x CH₂CO-Lev), 2.62-2.43 (m, 4 H, 2 x CH₂CH₂CO-Lev), 2.15, 2.14, 2.11, 2.10, 2.04, 1.99, 1.96 (7 x s, 7 x 3 H, 5 x Ac, 2 x CH₃-Lev); **¹³C NMR** (CDCl₃, 101 MHz) δ 206.33, 206.13 (2 x 90

COCH₃-Lev), 172.77, 171.33 (2 x COO), 171.14 (CONH), 170.63, 170.51, 170.40, 170.20 (4 x COOAc), 116.43 (CN), 101.09 (C-1'), 82.86 (C-1), 77.0 (C-5), 75.84 (C-4), 73.50 (C-3), 70.8 (C-3'), 70.8 (C-5'), 69.24 (C-2'), 66.82 (C-4'), 62.08 (C-6), 60.90 (C-6'), 52.86 (C-2'), 37.87, 37.68 (2 x CH₂COCH₃-Lev), 29.77, 29.75 (2 x COCH₃-Lev), 28.08, 27.77 (2 x CH₂COO), 23.21 (NHAc), 21.00, 20.79, 20.70, 20.62 (4 x COAc), 14.62 (CH₂CN); **MS** (ESI⁺) m/z = 803.81 [M+H]; [α]_D²⁰ = -117.3 (c = 1, CHCl₃); **IR** (ATR) 2218.52 (CN), 1715.33, 1722.74 (CO), 1699.29 (CONH).

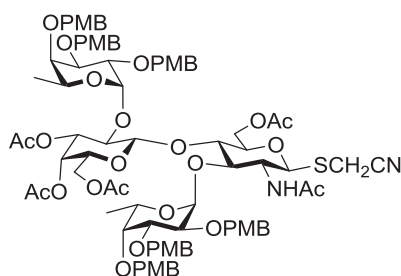
2-N-acetamido-6-O-acetyl-4-O-[3,4,6-tri-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.31**



2.30 (710 mg, 0.88 mmol, 1 eq.) was dried under vacuum for 30 min, then MeOH:DCM (dry, 1:1, 16 mL) and hydrazine acetate (405 mg, 4.4 mmol, 5 eq.) were added. The reaction was stirred at r.t. for 4 h before being quenched with acetylacetone (560 μ L) and stirred at r.t. for 30 min. The solvent was evaporated and the residue purified by flash column chromatography on silica (EtOH:CHCl₃, 0.25:10 to 0.5:10) to afford the title compound **2.31** as a light yellow solid (314.5 mg, 0.52 mmol, 59%). **TLC** R_f = 0.05 (EtOH:CHCl₃, 0.5:10); **¹H NMR** (CDCl₃, 250 MHz) δ 6.41 (d, *J* = 8.6 Hz, 1 H, NH), 5.35 (d, *J* = 3.2 Hz, 1 H, H-4'), 4.92 (dd, *J* = 10.2, 3.3 Hz, 1 H, H-3'), 4.76 (d, *J* = 10.1 Hz, 1 H, H-1), 4.57 (m, 1 H, H-6a),

4.47 (d, $J = 7.9$ Hz, 1 H, H-1'), 4.21 (dd, $J = 12.2, 5.2$ Hz, 1 H, H-6b), 4.13-4.01 (m, 3 H, H-6', H-5'), 3.96-3.86 (m, 1 H, H-2), 3.80 (dd, $J = 13.0, 4.8$ Hz, 1 H, H-2'), 3.68 (dt, $J = 11.0, 5.5$ Hz, 3 H, H-3, CH₂aCN, H-5), 3.60-3.51 (m, 1 H, H-4), 3.32 (d, $J = 17.1$ Hz, 1 H, SCH₂bCN), 2.13, 2.09, 2.03, 2.02, 2.01 (5 x s, 5 x 3 H, 5 x Ac); **¹³C NMR** (CDCl₃, 63 MHz) δ 171.64 (CO-OAc), 171.23 (CO-NHAc), 170.73, 170.47, 170.14 (3 x CO-OAc), 116.52 (CN), 104.30 (C-1'), 82.43 (C-1), 81.67 (C-4), 77.61 (C-5), 74.08 (C-3), 72.51 (C-3'), 71.69 (C-5'), 68.80 (C-2'), 67.16 (C-4'), 63.31 (C-6), 61.86 (C-6'), 54.35 (C-2), 23.46 (NHAc), 21.12, 20.82, 20.74, 20.70 (4 x OAc), 14.57 (CH₂CN); **MS** (ESI⁺) $m/z = 607.61$ [M+H]⁺; $[\alpha]_D^{20} = -28.3$ (c = 1, CHCl₃); **IR** (ATR) 3398.71 (OH), 2217.44 (CN), 1715.96 (CO), 1699.29 (CONH).

2-N-acetamido-6-O-acetyl-3-O-[2,3,4-tri-O-paramethoxybenzyl-6-deoxy- α -L-galactopyranosyl]-4-O-[2-(2,3,4-tri-O-paramethoxybenzyl-6-deoxy- α -L-galactopyranosyl)-3,4,6-tri-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside 2.32

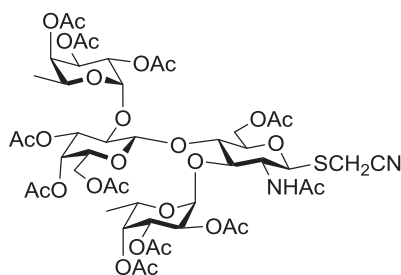


2.31 (294 mg, 450 μ mol, 1.0 eq.) and **2.20** (1022 mg, 1800 μ mol, 4.0 eq.), co-evaporated with toluene and dried under high vacuum for 1 h, were dissolved in

DCM:DMF (dry, 1:1, 2 mL) in a flask containing MS (4 Å, 450 mg), and stirred 1 h at r.t. before copper(II) bromide (201 mg, 900 µmol, 2.0 eq.) and tetrabutylammonium bromide (306 mg, 950 µmol, 2.1 eq.) were added. The mixture was stirred at r.t. for 16 h with light exclusion. The crude mixture was filtered over celite and washed with EtOAc (200 mL). The filtrate was washed with brine (150 mL) and NH₄Cl satd. solution (pH 9, 150 mL). The organic layer was dried over MgSO₄, filtered, concentrated under vacuum and purified by flash column chromatography on silica (EtOAc:PE, 1:2 to 1:0) to afford the title compound **2.32** as a white foam (521.2 mg, 0.32 mmol, 72%). **TLC** R_f = 0.95 (EtOH:CHCl₃, 0.5:10); **¹H NMR** (CDCl₃, 400 MHz) δ 7.25 (dddd, *J* = 20.9, 16.6, 11.8, 7.0 Hz, 12 H, PMB-Ar), 6.83 (qdd, *J* = 8.7, 4.0, 2.2 Hz, 12 H, PMB-Ar), 6.33 (d, *J* = 6.8 Hz, 1 H, NH), 5.31 (t, *J* = 4.0 Hz, 1 H, H-4'), 5.11 (d, *J* = 3.7 Hz, 1 H, H-1''), 5.08-5.02 (m, 3 H, H-3', H-1, H-1''), 4.89-4.77 (m, 4 H, 2 x CH₂-PMB), 4.63 (qdd, *J* = 21.7, 11.0, 6.5 Hz, 9 H, 4 x CH₂-PMB, H-6a), 4.45 (d, *J* = 7.9 Hz, 1 H, H-1'), 4.40 (d, *J* = 6.6 Hz, 1 H, H-5'''), 4.30-4.23 (m, 1 H, H-6b), 4.13-4.04 (m, 5 H, H-3, H-6', H-5'', H-2'''), 3.97 (dd, *J* = 10.2, 3.7 Hz, 1 H, H-2''), 3.91-3.86 (m, 2 H, H-4, H-3'''), 3.80-3.73 (m, 20 H, H-2', H-5', 6 x OMe), 3.71 (d, *J* = 3.0 Hz, 1 H, H-3''), 3.68 (m, 1 H, H-4''), 3.66 (m, 1 H, H-4'''), 3.62-3.55 (m, 2 H, H-5, H-2), 3.48 (d, *J* = 17.1 Hz, 1 H, SCH₂aCN), 3.22 (d, *J* = 17.0 Hz, 1 H, SCH₂bCN), 2.06, 2.01, 1.96, 1.78, 1.67 (5 x s, 5 x 3H, 5 x Ac), 1.15 (2 x d, *J* = 6.4 Hz, 2 x 3 H, H-6'', H-6'''); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.83 (CO-NHAc), 170.38, 170.30, 170.04, 169.92 (4 x CO-OAc), 159.63, 159.37, 159.29, 159.26, 159.25, 159.24 (6 x C_{Ar}-OMe), 130.98, 130.91, 130.79, 130.75, 130.73, 130.68 (6 x C_{Ar}-CH₂O), 130.10, 129.98, 129.82, 129.48, 129.14, 128.95 (12 x CH_{Ar}-CCH₂OR), 116.74 (CN), 114.23, 113.96, 113.94, 113.76, 113.75, 113.74 (12 x CH_{Ar}-COMe), 101.22 (C-1'), 99.18 (C-1''),

98.68 (C-1'''), 81.56 (C-1), 80.16 (C-3'''), 79.22 (C-3''), 77.5 (C-5), 77.2 (C-4''), 76.62, 76.54 (C-3, C-2), 76.33 (C-4'''), 75.69 (C-2''), 74.62, 74.58 (2 x CH_{Ar}CCH₂OR), 74.40 (C-4), 74.33, 74.13 (2 x CH_{Ar}CCH₂OR), 73.52 (C-2'), 73.22 (C-4''), 72.82 (C-3'), 72.57, 72.20 (2 x CH_{Ar}CCH₂OR), 70.89 (C-5'), 67.66 (C-5''), 67.30 (C-4'), 67.01 (C-5'''), 62.50 (C-6), 60.77 (C-6'), 56.50 (C-2), 55.48, 55.42, 55.38 (6 x OMe), 23.10 (NHAc), 21.03, 20.92, 20.80, 20.68 (4 x OAc), 17.07 (C-6'''), 16.75 (C-6''), 14.38 (CH₂CN); **MS** (ESI⁺) *m/z* = 1620.78 [M+H]⁺; [α]_D²⁰ = -49.1 (c = 1, CHCl₃); **IR** (ATR) 2218.19 (CN), 1715.37 (CO), 1698.09 (CONH).

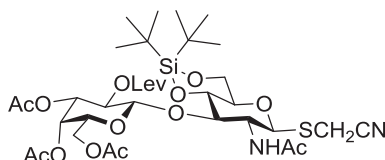
2-N-acetamido-6-O-acetyl-3-O-[2,3,4-tri-O-acetyl-6-deoxy-α-L-galactopyranosyl]-4-O-[2-(2,3,4-tri-O-acetyl-6-deoxy-α-L-galactopyranosyl)-3,4,6-tri-O-acetyl-β-D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy-β-D-glucopyranoside 2.33



To **2.32** (469.7 mg, 0.29 mmol, 1.0 eq.) dissolved in ACN:H₂O (9:1, 10 mL) was added CAN (2.099 g, 3.82 mmol, 13.2 eq.). The solution was stirred at r.t. for 4 h, then Ac₂O (23 mL), pyridine (23 mL) and DMAP (3.5 mg, 0.029 mmol, 0.1 eq.) were added. The reaction was stirred at r.t. for 16 h, then quenched with MeOH. The solvents were evaporated, the crude mixture was recovered with EtOAc (100

mL), washed with H₂O (100 mL), NaHCO₃ (100 mL) and brine (100 mL), dried over MgSO₄, evaporated and purified by flash column chromatography on silica (EtOAc:PE, 1:10 to 1:0) to afford the title compound **2.33** as a white powder (238 mg, 0.21 mmol, 71%). **TLC** R_f = 0.75 (EtOH:CHCl₃, 0.5:10); **¹H NMR** (CDCl₃, 400 MHz) δ 5.97 (d, *J* = 9.7 Hz, 1 H, NH), 5.46-5.30 (m, 5 H, H-1'', H-4'', H-4''', H-4', H-1'''), 5.19 (dd, *J* = 10.9, 3.2 Hz, 1 H, H-3'''), 5.09 (dd, *J* = 10.3, 8.3 Hz, 1 H, H-3''), 5.09-4.95 (m, 3 H, H-3', H-2'', H-2'''), 4.92 (dd, *J* = 13.2, 6.7 Hz, 1 H, H-5''), 4.70 (dd, *J* = 6.1, 2.6 Hz, 1 H, H-1), 4.68 (d, *J* = 12.0 Hz, 1 H, H-6a), 4.55-4.46 (m, 2 H, H-1', H-6'a), 4.39 (dd, *J* = 12.9, 6.7 Hz, 1 H, H-6'b), 4.33-4.24 (m, 1 H, H-5'''), 4.17 (dd, *J* = 12.2, 4.5 Hz, 1 H, H-6b), 3.98-3.81 (m, 4 H, H-3, H-2, H-2', H-5'), 3.60 (d, *J* = 17.4 Hz, 1 H, CH₂aCN), 3.55 (ddd, *J* = 5.7, 3.3, 2.0 Hz, 1 H, H-5), 3.28 (d, *J* = 17.2 Hz, 1 H, CH₂bCN), 2.18, 2.14, 2.13, 2.12, 2.10, 2.09, 2.06, 2.06, 2.03, 1.98, 1.97 (11 x s, 11x 3 H, 11 x Ac), 1.22 (d, *J* = 3.2 Hz, 3 H, H-6''), 1.20 (d, *J* = 7.0 Hz, 3 H, H-6'''); **¹³C NMR** (CDCl₃, 101 MHz) δ 171.68, 171.35, 170.93, 170.87, 170.79, 170.72, 170.61, 170.57, 170.26, 170.00, 169.71 (11 x CO), 117.49 (CN), 100.79 (C-1'), 96.49 (C-1'''), 96.00 (C-1''), 82.28 (C-1), 77.27 (C-5), 75.54 (C-3), 74.17 (C-4), 73.59 (C-3'), 72.97 (C-2'), 71.49 (C-4''), 71.22 (C-4'''), 71.09 (C-5'), 68.25 (C-2''), 68.07 (C-2'''), 68.00 (C-3''), 67.40 (C-3'''), 66.96 (C-4'), 65.27 (C-5'''), 64.30 (C-5''), 61.96 (C-6), 60.78 (C-6'), 55.71 (C-2), 23.57 (NHAc), 21.21, 21.01, 20.97, 20.88, 20.87, 20.85, 20.83, 20.81, 20.76, 20.75 (10 x OAc), 16.06 (C-6''), 15.99 (C-6'), 14.34 (CH₂CN); **MS** (ESI⁺) *m/z* = 1149.34 [M⁺]; [α]_D²⁰ = -147.4 (c = 1, CHCl₃); **IR** (ATR) 2208.53 (CN), 1744.21 (CO).

2-N-acetamido-4,6-O-di-*tert*-butyl-dimethylsilylidene-3-O-[2-O-levulinoyl-3,4,6-tri-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.34**

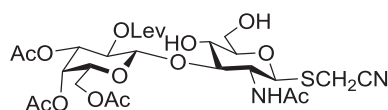


2.6 (1.0 g, 2.4 mmol, 1 eq.) and **2.15** (1.68 g, 3.6 mmol, 1.5 eq.) were co-evaporated with toluene and dried under high vacuum for 1 h. The mixture was dissolved in DCM (dry, 24 mL), added to a flame-dried flask with MS (4 Å, 2.4 g) and stirred at r.t. for 1 h before addition of DTBMP (493 mg, 2.4 mmol, 1 eq.) and AgOTf (1.418 g, 5.52 mmol, 2.3 eq.). The mixture was stirred at r.t. for 2.5 h in the dark. The reaction mixture was filtered over Celite and rinsed with DCM. After removal of the solvent, the residue was purified by flash column chromatography on silica (EtOH:CHCl₃, 0:10 to 2:10) to afford the title compound **2.34** as a white amorphous solid (1.2 g, 1.49 mmol, 63%). **TLC** R_f = 0.45 (EtOH:CHCl₃, 1:10); **¹H NMR** (CDCl₃, 400 MHz) δ 6.86 (d, *J* = 7.8 Hz, 1 H, NH), 5.46-5.25 (m, 1 H, H-4'), 5.11-4.84 (m, 4 H, H-1, H-2', H-1', H-3'), 4.25 (dd, *J* = 7.1, 3.9 Hz, 1 H, H-6a), 4.17-3.95 (m, 4 H, H-3, H-6', H-4), 3.92-3.78 (m, 2 H, H-6b, H-5'), 3.58-3.42 (m, 3 H, H-2, SCH₂aCN, H-5), 3.30 (d, *J* = 17.0 Hz, 1 H, SCH₂bCN), 2.79-2.68 (m, 2 H, Lev), 2.56-2.49 (m, 1 H, Lev), 2.19, 2.14, 2.08, 2.02, 1.98, 1.94, 1.91 (7 x s, 7 x 3 H, 4 x Ac, 3 x CH₃-Lev), 1.00, 0.94 (2 x s, 2 x 9 H, 2 x ^tBu); **¹³C NMR** (CDCl₃, 101 MHz) δ 209.40 (CO-Lev), 171.45 (COO-Lev), 171.10, 170.44, 170.40, 170.36 (4 x CO-OAc), 116.65 (CN), 100.43 (C-1'), 82.71 (C-1), 80.40 (C-3), 77.28

(C-4), 74.95 (C-5), 71.76 (C-2'), 71.26 (C-3'), 70.69 (C-5'), 66.83 (C-4'), 66.29 (C-6), 61.25 (C-6'), 54.6 (C-2), 37.93 (**CH**₂COCH₃-Lev), 30.10 (CH₂CO**CH**₃-Lev), 28.03 (OCO**CH**₂-Lev), 27.44, 27.18 (2 x C(**CH**₃)₃), 23.46 (OAc), 22.74 (**C**(CH₃)₃), 20.85, 20.77, 20.75 (3 x OAc), 20.09 (**C**(CH₃)₃), 14.36 (**CH**₂CN); **MS** (ESI⁺) *m/z* = 803.97 [M+H]⁺; [α]_D²⁰ = -41.4 (c = 1, CHCl₃); **IR** (ATR) 2217.48 (CN), 1717.78 (CO), 1699.00 (CONH).

2-N-acetamido-3-O-[2-O-levulinoyl-3,4,6-tri-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside

2.35

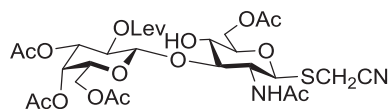


To **2.34** (1.2 g, 1.5 mmol, 1 eq.) in THF (dry, 15 mL) was added triethylamine trihydrofluoride (513.5 μ L, 3.1 mmol, 2.1 eq.). After 3 h the solvent was evaporated and the residue purified by flash column chromatography on silica (EtOAc:MeOH, 10:1 to 10:2), to afford the title compound **2.35** as a white amorphous solid (566.5 mg, 0.85 mmol, 57%). **TLC** *R*_f = 0.45 (EtOAc:MeOH, 10:0.5); **¹H NMR** (CDCl₃, 400 MHz) δ 7.76 (d, *J* = 6.8 Hz, 1 H, NH), 5.51 (d, *J* = 10.2 Hz, 1 H, H-1), 5.36 (d, *J* = 2.9 Hz, 1 H, H-4'), 5.21 (dd, *J* = 10.5, 8.0 Hz, 1 H, H-2'), 5.02 (dd, *J* = 10.6, 3.4 Hz, 1 H, H-3'), 4.50 (d, *J* = 8.0 Hz, 1 H, H-1'), 4.40 (t, *J* = 8.9 Hz, 1 H, H-3), 4.15-4.01 (m, 3 H, H-6', H-5'), 3.94 (dd, *J* = 12.0, 2.7 Hz, 1 H, H-6a), 3.75 (dd, *J* = 12.0, 5.6 Hz, 1 H, H-6b), 3.59-3.53 (m, 2 H, H-5,

CH₂aCN), 3.48 (dd, $J = 10.8, 6.9$ Hz, 1 H, H-4), 3.40 (d, $J = 17.2$ Hz, 1 H, CH₂bCN), 3.20-2.96 (m, 2 H, Lev), 2.68-2.56 (m, 2 H, Lev), 2.42-2.31 (m, 1 H, Lev), 2.25, 2.13, 2.05, 2.02, 1.96 (5 x s, 5 x 3 H, 4 x Ac, CH₃-Lev); **¹³C NMR** (CDCl₃, 101 MHz) δ 209.73 (CO-Lev), 171.45 (COO-Lev), 171.13, 170.63, 170.19, 170.05 (4 x CO-OAc), 117.64 (CN), 101.54 (C-1'), 83.40 (C-3), 81.67 (C-1), 80.24 (C-5), 71.23 (C-5'), 70.59 (C-3'), 70.03 (C-4), 69.23 (C-2'), 67.03 (C-4'), 63.09 (C-6), 61.55 (C-6'), 56.51 (C-2), 37.75 (CH₂COCH₃-Lev), 30.23 (CH₂COCH₃-Lev), 27.61 (OCOCH₂-Lev), 23.69, 21.17, 20.69, 20.64 (4 x Ac), 15.06 (SCH₂CN); **MS** (ESI⁺) $m/z = 663.67$ [M+H]; $[\alpha]_D^{20} = +38.3$ (c = 1, CHCl₃); **IR** (ATR) 3401.98 (OH), 2217.94 (CN), 1717.67 (CO), 1697.99 (CONH).

2-N-acetamido-6-O-acetyl-3-O-[2-O-levulinoyl-3,4,6-tri-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside

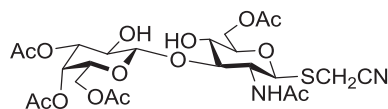
2.36



To **2.35** (566.5 mg, 0.85 mmol, 1 eq.) in DCM (dry, 12 mL) and pyridine (dry, 275 μ L, 3.4 mmol, 4 eq.) under argon at -78 °C was added AcCl (181 μ L, 2.55 mmol, 3 eq.). After 1 h MeOH was added to quench the reaction and stirred at -78 °C for a few min. The crude was evaporated, recovered with DCM (150 mL) and washed with HCl 1N (100 mL), NaHCO₃ (100 mL), brine (100 mL) and dried over MgSO₄. Evaporation of the solvent afforded the title compound **2.36** as a white amorphous solid (559.5 mg, 0.79 mmol, 93%). **TLC** $R_f = 0.6$ (EtOAc:MeOH, 10:1);

¹H NMR (CDCl₃, 400 MHz) δ 7.78-7.73 (m, 1 H, NH), 5.37 (m, 2 H, H-1, H-4'), 5.25-5.18 (m, 1 H, H-2'), 5.09-5.01 (m, 1 H, H-3'), 4.55 (d, *J* = 8.0 Hz, 1 H, H-1'), 4.43 (dd, *J* = 12.0, 1.7 Hz, 1 H, H-3), 4.35 (t, *J* = 8.9 Hz, 1 H, H-6a), 4.23 (dd, *J* = 12.1, 5.8 Hz, 1 H, H-6b), 4.14-4.04 (m, 3 H, H-6', H-5'), 3.63 (m, 1 H, OH), 3.58-3.54 (m, 1 H, H-5), 3.50 (dd, *J* = 9.7, 8.5 Hz, 1 H, H-4), 3.36 (dd, *J* = 17.0, 7.0 Hz, 1 H, CH₂aCN), 3.24 (d, *J* = 7.8 Hz, 1 H, CH₂bCN), 3.08-2.97 (m, 1 H, Lev), 2.71-2.56 (m, 2 H, Lev), 2.43-2.32 (m, 1 H, Lev), 2.24, 2.12, 2.07, 2.04, 2.02, 1.97 (6 x s, 6 x 3H, 5 x Ac, CH₃-Lev); **¹³C NMR** (CDCl₃, 101 MHz) δ 209.63 (CO-Lev), 171.41 (COO-Lev), 171.21, 171.00, 170.61, 170.15, 170.00 (5 x CO-Ac), 116.42 (CN), 101.58 (C-1'), 83.34 (C-3), 80.29 (C-1), 78.17 (C-5), 71.24 (C-5'), 70.58 (C-3'), 69.37 (C-4), 69.18 (C-2'), 67.03 (C-4'), 63.41 (C-6), 61.53 (C-6'), 56.72 (C-2), 37.74 (CH₂COCH₃-Lev), 30.17 (CH₂COCH₃-Lev), 27.60 (OCOCH₂-Lev), 23.67, 21.03, 20.97, 20.68, 20.62 (5 x Ac), 14.26 (CH₂CN); **MS** (ESI⁺) *m/z* = 705.71 [M+H]; [α]_D²⁰ = +12.0 (*c* = 1, CHCl₃); **IR** (ATR) 3401.57 (OH), 2216.90 (CN), 1715.52 (CO), 1698.43 (CONH).

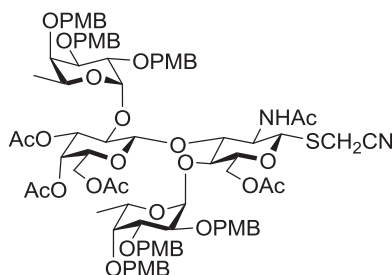
2-N-acetamido-6-O-acetyl-3-O-[3,4,6-tri-O-acetyl-β-D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy-β-D-glucopyranoside 2.37



2.36 (559.5 mg, 0.79 mmol, 1 eq.) was dried under vacuum for 30 min, then MeOH:DCM (dry, 1:1, 16 mL) and hydrazine acetate (87.7 mg, 0.95 mmol, 1.2

eq.) were added and the mixture stirred at r.t.. After 4 h the solvent was evaporated and the residue purified by flash column chromatography on silica (EtOH:CHCl₃, 0.2:10) to afford the title compound **2.37** as a white solid (195 mg, 0.32 mmol, 41%). **TLC** R_f = 0.55 (EtOAc:MeOH, 10:1); **¹H NMR** (CD₃CN, 400 MHz) δ 6.77 (d, *J* = 9.0 Hz, 1 H, NH), 5.31 (d, *J* = 3.2 Hz, 1 H, H-4'), 4.91 (dd, *J* = 10.1, 3.5 Hz, 1 H, H-3'), 4.81 (d, *J* = 10.5 Hz, 1 H, H-1), 4.51 (d, *J* = 7.9 Hz, 1 H, H-1'), 4.39 (dd, *J* = 12.0, 2.0 Hz, 1 H, H-6a), 4.24 (dd, *J* = 12.1, 6.1 Hz, 2 H, H-6b, OH), 4.18-4.06 (m, 3 H, H-5', H-6'), 3.96-3.82 (m, 2 H, H-2, OH), 3.69 (qd, *J* = 9.3, 4.2 Hz, 3 H, H-3, H-2', CH₂aCN), 3.60 (ddd, *J* = 9.6, 6.0, 2.0 Hz, 1 H, H-5), 3.54-3.44 (m, 2 H, CH₂bCN, H-4), 2.12, 2.06, 2.03, 2.00, 1.91 (5 x s, 5 x 3 H, 5 x Ac); **¹³C NMR** (CD₃CN, 101 MHz) δ 171.93, 171.51, 171.21, 171.03 (5 x CO-Ac), 118.3 (CN), 105.04 (C-1'), 87.61 (C-3), 83.67 (C-1), 78.62 (C-5), 73.54 (C-3'), 71.99 (C-5'), 70.03 (C-4), 69.35 (C-2'), 68.19 (C-4'), 63.92 (C-6), 62.59 (C-6'), 53.90 (C-2), 23.23, 20.95, 20.90, 20.77, 20.74 (5 x Ac), 15.37 (CH₂CN); **MS** (ESI⁺) *m/z* = 607.61 [M+H]; [α]_D²⁰ = -29.9 (c = 1, acetone); **IR** (ATR) 3400.07 (OH), 2215.89 (CN), 1716.66 (CO), 1697.38 (CONH).

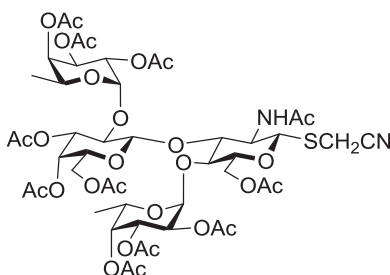
2-N-acetamido-6-O-acetyl-4-O-[2,3,4-tri-O-paramethoxybenzyl-6-deoxy- α -L-galactopyranosyl]-3-O-[2-(2,3,4-tri-O-paramethoxybenzyl-6-deoxy- α -L-galactopyranosyl)-3,4,6-tri-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.38**



2.37 (195 mg, 320 μ mol, 1.0 eq.) and **2.20** (728 mg, 1280 μ mol, 4.0 eq.), co-evaporated with toluene and dried under high vacuum for 1 h, were dissolved in DCM:DMF (dry, 1:1, 2 mL) with MS (4 Å, 320 mg) and stirred 1 h at r.t. before copper(II) bromide (286 mg, 1280 μ mol, 4.0 eq.) and tetra-butylammonium bromide (432 mg, 1340 μ mol, 4.2 eq.) were added. The mixture was stirred at r.t. for 16 h with light exclusion. The crude mixture was filtered over celite and washed with DCM (200 mL). The filtrate was evaporated, recovered with EtOAc (100 mL), washed with NH_4Cl satd. solution (100 mL, pH 9) and brine (100 mL). The organic layer was dried over MgSO_4 , filtered, concentrated under vacuum and purified by flash column chromatography on silica (EtOAc:PE, 2:8 to 1:0) to afford the title compound **2.38** as a white solid (208.8 mg, 0.13 mmol, 40%). **TLC** R_f = 0.2 (EtOH: CHCl_3 , 10:0.05); **^1H NMR** (CDCl_3 , 101 MHz) δ 7.42-7.17 (m, 12 H, PMB-Ar), 6.96-6.80 (m, 12 H, PMB-Ar), 6.60 (d, J = 8.1 Hz, 1 H, NH), 5.31 (d, J = 3.4 Hz, 1 H, H-4'), 5.24 (d, J = 3.7 Hz, 1 H, H-1''), 5.11 (m, 1 H, H-1), 4.99

(dd, $J = 10.0, 3.4$ Hz, 1 H, H-3'), 4.90 (dd, $J = 13.8, 11.5$ Hz, 2 H, CH₂PMB), 4.83 (d, $J = 3.1$ Hz, 1 H, H-1'''), 4.81 (m, 1 H, CH₂PMB), 4.74 (dd, $J = 7.1, 3.9$ Hz, 4 H, 2 x CH₂PMB), 4.72-4.64 (m, 3 H, CH₂PMB), 4.63-4.59 (m, 3 H, CH₂PMB, H-1'), 4.58-4.52 (m, 2 H, H-5''', H-6a), 4.38 (d, $J = 10.2$ Hz, 1 H, H-6b), 4.22 (br, 1 H, H-3), 4.14 (tdd, $J = 10.3, 6.5, 3.6$ Hz, 3 H, H-6', H-2''), 4.06-3.95 (m, 3 H, H-5'', H-2''', H-2'), 3.90-3.79 (m, 22 H, H-5', H-3'', H-4'', H-3''', 6 x OMe), 3.72 (d, $J = 4.5$ Hz, 2 H, H-4, H-5), 3.66 (m, 2 H, H-4''', H-2), 3.52 (d, $J = 17.0$ Hz, 1 H, CH₂aCN), 3.29 (d, $J = 17.0$ Hz, 1 H, CH₂bCN), 2.08, 2.04, 1.89, 1.85, 1.80 (5 x s, 5 x 3 H, 5 x Ac), 1.17 (dd, $J = 10.2, 6.5$ Hz, 6 H, H-6'', H-6'''); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.66, 170.51, 170.29, 169.98, 169.92 (5 x CO-Ac), 159.40, 159.23, 159.17, 159.11, 159.06, 159.03 (6 x C_{Ar}-OMe), 130.94, 130.83, 130.76, 130.65, 130.61 (5 x C_{Ar}-CH₂OR), 130.59, 130.44 (4 x CH_{Ar}-CCH₂OR), 130.09 (6 x C_{Ar}-CH₂OR), 130.03, 129.07, 129.03, 128.71 (8 x CH_{Ar}-CCH₂OR), 116.60 (CN), 113.85, 113.82, 113.74, 113.61, 113.57 (12 x CH_{Ar}-COMe), 102.36 (C-1''), 98.00 (C-1'''), 97.83 (C-1'), 81.40 (C-1), 80.24 (C-3'''), 79.28 (C-3''), 77.43 (C-5), 77.12 (C-3), 76.80 (C-4''), 76.52 (C-4'''), 75.90 (C-2''), 75.54 (C-2'''), 74.71, 74.39 (4 x CH₂OR-PMB), 73.65 (C-3'), 73.39 (C-4), 73.17, 72.87 (2 x CH₂OR-PMB), 72.19 (C-2'), 70.15 (C-5'), 67.45 (C-5''), 67.01 (C-4'), 66.64 (C-5'''), 61.97 (C-6), 60.23 (C-6'), 55.27 (C-2), 55.25 (6 x OMe), 23.24 (NHAc), 20.92, 20.75, 20.70, 20.45 (4 x OAc), 16.86 (C-6'''), 16.47 (C-6''), 14.68 (CH₂CN); **MS** (ESI⁺) m/z = 1620.78 [M+H]; $[\alpha]_D^{20} = -42.3$ ($c = 1$, CHCl₃); **IR** (ATR) 2215.79 (CN), 1716.05 (CO), 1697.70 (CONH).

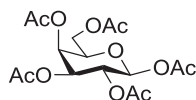
2-N-acetamido-6-O-acetyl-4-O-[2,3,4-tri-O-acetyl-6-deoxy- α -L-galactopyranosyl]-3-O-[2-(2,3,4-tri-O-acetyl-6-deoxy- α -L-galactopyranosyl)-3,4,6-tri-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.39**



To **2.38** (208.8 mg, 0.129 mmol, 1.0 eq.) dissolved in ACN:H₂O (9:1, 10 mL) was added CAN (954 mg, 1.74 mmol, 13.5 eq.). The solution was stirred at r.t. for 4 h, then Ac₂O (13 mL), pyridine (13 mL) and DMAP (1.6 mg, 0.0129 mmol, 0.1 eq.) were added. The reaction was stirred at r.t. for 16 h, then quenched with MeOH. Solvents were evaporated, the crude mixture was recovered with EtOAc (150 mL), washed with H₂O (150 mL), NaHCO₃ (150 mL), brine (150 mL), dried over MgSO₄, evaporated and purified by flash column chromatography on silica (EtOAc:PE, 1:10 to 1:0) to afford the title compound **2.39** as a white powder (112 mg, 97.3 μ mol, 76%). **TLC** R_f = 0.5 (EtOH:CHCl₃, 10:1); **¹H NMR** (CDCl₃, 400 MHz) δ 6.70 (d, *J* = 8.7 Hz, 1 H, NH), 5.49 (dd, *J* = 10.9, 3.3 Hz, 1 H, H-4'), 5.38-5.25 (m, 5 H, H-2'', H-1'', H-4'', H-4''', H-2'''), 5.16 (dd, *J* = 11.1, 3.1 Hz, 1 H, H-3'''), 5.03-4.89 (m, 5 H, H-3', H-5''', H-3'', H-1, H-1''), 4.72 (d, *J* = 7.9 Hz, 1 H, H-1'), 4.60-4.43 (m, 5 H, H-5'', H-6'a, H-4, H-3, H-6a), 4.38 (m, 1 H, H-6'b), 4.21 (dd, *J* = 11.2, 8.6 Hz, 1 H, H-6b), 4.10 (q, *J* = 7.1 Hz, 1 H, H-2'), 3.98-3.83 (m, 2

H, H-2, H-5), 3.56 (d, $J = 17.2$ Hz, 1 H, CH₂aCN), 3.37 (d, $J = 17.2$ Hz, 1 H, CH₂bCN), 2.15, 2.14, 2.12, 2.11, 2.09, 2.07, 2.05, 1.99, 1.96, 1.94, 1.94 (11 x s, 11 x 3 H, 11 x Ac), 1.28-1.13 (m, 6 H, H-6'', H-6'''); **¹³C NMR** (CDCl₃, 101 MHz) δ 171.52 (CO-NHAc), 170.84, 170.81, 170.61, 170.53, 170.47, 170.38, 170.33, 170.26, 170.19, 170.07 (10 x CO-OAc), 116.84 (CN), 101.33 (C-1'), 96.73 (C-1), 96.71 (C-1'''), 96.29 (C-1''), 77.91 (C-5), 76.38 (C-2'), 73.82 (C-2), 73.77 (C-3), 72.31 (C-4), 71.72 (C-4''), 71.45 (C-4'''), 70.35 (C-5'), 68.71 (C-3'), 67.99 (C-3'''), 67.36 (C-3''), 67.13 (C-4'), 66.83 (C-2''), 65.23 (C-2'''), 64.80 (C-5''), 64.46 (C-5'''), 61.43 (C-6), 60.45 (C-6'), 22.98 (NHAc), 21.10, 20.86, 20.80, 20.73, 20.65 (10 x OAc), 15.93 (C-6'''), 15.50 (C-6''), 14.20 (CH₂CN); **MS** (ESI⁺) $m/z = 1152.28$ [M+H]⁺; $[\alpha]_D^{20} = -127.3$ (c = 1, CHCl₃); **IR** (ATR) 2209.47 (CN), 1744.16 (CO).

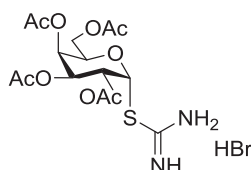
1,2,3,4,6-penta-O-acetyl - β -D-galactopyranoside **2.40**^[18]



To acetic anhydride (200 mL) was added D-galactopyranoside (5 g, 27.25 mmol, 1 eq.) under nitrogen atmosphere and strong stirring, followed by pyridine (100 mL) and DMAP (678 mg, 5.55 mmol, 0.2 eq). The mixture was stirred for 1 h at r.t., before being slowly quenched with MeOH (200 mL), poured into ice (200 g) and water (50 mL) and recovered with EtOAc (250 mL). The organic layer was separated and washed with HCl 1M (250 mL), NaHCO₃ satd. (250 mL), brine (250 mL) and dried over MgSO₄. Finally, the solution was concentrated under

vacuum to afford the product **2.40** as a yellow oil (10.82 g, 27.2 mmol, quant). **TLC** R_f = 0.45 (EtOAc:PE, 1:1); **^1H NMR** (CDCl_3 , 400 MHz) δ 6.28 (d, J = 2.7 Hz, 1 H, H-1), 5.41 (d, J = 1.3 Hz, 1 H, H-4), 5.29-5.15 (m, 2 H, H-2, H-3), 4.27 (dd, J = 7.1, 6.2 Hz, 1 H, H-5), 4.09-3.93 (m, 2 H, H-6), 2.07, 2.06, 1.95, 1.93, 1.91 (5 x s, 5 x 3 H, 5 x Ac); **^{13}C NMR** (CDCl_3 , 101 MHz) δ 170.67, 170.48, 170.34, 170.01, 169.95 (5 x CO), 80.63 (C-1), 77.27 (C-5), 75.39 (C-3), 66.12 (C-4), 65.91 (C-2), 61.37 (C-6), 20.89, 20.74, 20.57, 20.33, 20.19 (5 x OAc); **MS** (ESI^+) m/z = 391.35 $[\text{M}+\text{H}]$; $[\alpha]_D^{20}$ = +115.3 (c = 1.01, CHCl_3); **IR** (ATR) 1792.12 (CO).

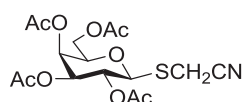
2,3,4,6-tetra-O-acetyl- α -D-galactopyranosyl thiouronium bromide **2.41**^[19]



To a solution of **2.10** (7.71 g, 18.76 mmol, 1 eq.) in acetone (60 mL) was added thiourea (3.14 g, 41.27 mmol, 2.2 eq.). The mixture was stirred for 2 h at reflux (80 °C). The reaction was cooled to r.t., filtered and the solid washed with cold EtOH, to afford the title compound **2.41** as a white solid (7.6 g, 18.7 mmol, quant.). **TLC** R_f = 0.0 (EtOAc:PE, 1:1); **^1H NMR** (DMSO-d_6 , 400 MHz) δ 9.40, 9.23 (2 x s, 4 H, NH_2 , NH_2Br), 5.68 (d, J = 3.9 Hz, 1 H, H-1), 5.25 (dd, J = 3.2, 1.1 Hz, 1 H, H-4), 5.37 (dd, J = 10.6, 3.3 Hz, 1 H, H-3), 4.99 (dd, J = 10.6, 4.0 Hz, 1 H, H-2), 4.43 (t, J = 6.6 Hz, 1 H, H-5), 4.17-4.03 (m, 2 H, H-6), 2.12, 2.10, 2.07, 2.02 (4 x s, 4 x 3 H, 4 x Ac); **^{13}C NMR** (DMSO-d_6 , 101 MHz) δ 170.20, 170.17, 169.99,

169.70 (4 x CO), 157.91 (CNH₂), 79.44 (C-1), 77.22 (C-5), 71.02 (C-3), 68.11 (C-4), 66.72 (C-2), 61.01 (C-6), 20.97, 20.76, 20.44, 20.32 (4 x OAc); **MS** (ESI⁺) *m/z* = 406.41 [M⁺]; [α]_D²⁰ = -26.0 (c = 1.03, H₂O); **IR** (ATR) 3233.53, 3178.43 (NH₂), 1743.11 (CO).

2,3,4,6-tetra-O-acetyl-1-thio-(S-cyanomethyl)- β -D-galactopyranoside **2.42**

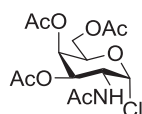


To a solution of **2.41** (5.7 g, 14.1 mmol, 1 eq.) in acetone:H₂O (1:1, 100 mL) was added K₂CO₃ (3.9 g, 28.14 mmol, 1.5 eq), Na₂S₂O₅ (9.6 g, 50.65 mmol, 2.7 eq) and chloroacetonitrile (6 mL, 93.8 mmol, 5 eq). The mixture was stirred for 12 h at r.t.. After this time the acetone was evaporated, and the aqueous layer extracted with DCM. The organic layer was washed with brine (125 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum and the product precipitated by adding water. The precipitate was washed with cold Et₂O and dried to afford the title compound **2.42** as a white solid (4.51 g, 11.15 mmol, 79%). **TLC** R_f = 0.75 (EtOAc:PE, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 5.33 (d, *J* = 2.7 Hz, 1 H, H-4), 5.08 (t, *J* = 9.8 Hz, 1 H, H-2), 5.01 (dd, *J* = 10.0, 3.3 Hz, 1 H, H-3), 4.64 (d, *J* = 9.7 Hz, 1 H, H-1), 4.06-3.91 (m, 3 H, H-6, H-5), 3.53 (d, *J* = 17.1 Hz, 1 H, CH_{2a}CN), 3.29 (d, *J* = 17.1 Hz, 1 H, CH_{2b}CN), 2.03, 1.95, 1.92, 1.85 (4 x s, 4 x 3 H, 4 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.30, 170.07, 169.79, 169.70 (4 x CO), 116.14 (CN), 82.14 (C-1), 74.73 (C-5), 71.32 (C-3), 67.16 (C-

4), 66.82 (C-2), 61.35 (C-6), 20.57, 20.52, 20.47, 20.44 (4 x OAc), 14.48 (CH₂CN); **MS** (ESI⁺) *m/z* = 404.41 [M+H]⁺; [α]_D²⁰ = -97.1 (c = 1.0, CHCl₃);^[20] **IR** (ATR) 2187.02 (CN), 1712.53 (CO).

2-N-acetamido-3,4,6-tri-O-acetyl-1-chloro-2-deoxy-α-D-galactopyranoside

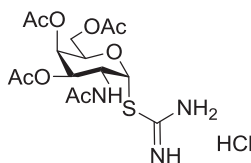
2.43^[21]



To AcCl (8 mL, 112 mmol, 9 eq.) was added N-acetyl galactosamine (2.75 g, 12.43 mmol, 1 eq.) under nitrogen atmosphere and strong stirring. The suspension was stirred for 48 h at r.t. Then, the reaction was recovered with DCM (150 mL) and poured into ice (50 g) and water (150 mL). The organic layer was separated and poured into a satd. solution of NaHCO₃ (200 mL). The mixture was stirred and then shaken into a separatory funnel until the production of gas ceased. The organic phase was then dried over MgSO₄ and evaporated under vacuum to afford the title compound **2.43** as a yellow oil (3.25 g, 8.88 mmol, 71%). **TLC** R_f = 0.8 (EtOAc, 100%); **¹H NMR** (CDCl₃, 200 MHz) δ 6.26 (d, *J* = 3.7 Hz, 1 H, H-4), 5.79 (d, *J* = 8.8 Hz, 1 H, H-3), 5.45 (dd, *J* = 3.1, 1.2 Hz, 1 H, H-1), 5.27 (dd, *J* = 11.3, 3.2 Hz, 1 H, H-2), 4.77 (ddd, *J* = 11.4, 8.9, 3.7 Hz, 1 H, H-5), 4.24-3.98 (m, 2 H, H-6), 2.15, 2.09, 2.05, 2.01 (4 x s, 4 x 3 H, 4 x Ac); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.89, 170.77, 170.59, 170.34 (4 x CO), 83.22 (C-1), 74.06 (C-5), 70.28 (C-3), 66.75 (C-4), 61.64 (C-6), 48.00 (C-2), 23.15 (NHAc), 20.91, 20.88, 20.43 (3 x OAc); **MS** (ESI⁺) *m/z* = 366.77 [M+H]⁺; [α]_D²⁰ = +56.2 (c = 1, CHCl₃); **IR** (ATR) 1748.75 (CO).

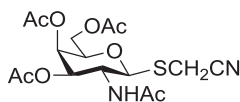
2-N-acetamido-3,4,6-tri-O-acetyl-2-deoxy- α -D-galactopyranosyl

thiouronium chloride **2.44**



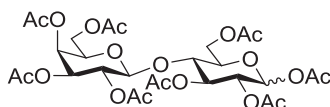
To a solution of **2.43** (3.25 g, 8.89 mmol, 1 eq.) in acetone (35 mL) was added thiourea (1.35 g, 17.78 mmol, 2 eq.). The mixture was stirred for 2 h at reflux (80 °C). Then, the reaction was cooled to r.t., filtered and the solid washed with cold EtOH to afford the title compound **2.44** as a white solid (2.269 g, 5.6 mmol, 63%). **TLC** R_f = 0.0 (EtOAc, 100%); **^1H NMR** (DMSO- d_6 , 400 MHz) δ 9.14, 9.39 (2 x s, 4 H, NH_2 , NH_2Cl), 8.39 (d, J = 9.0 Hz, 1 H, NH), 5.58 (d, J = 10.5 Hz, 1 H, H-1), 5.35 (d, J = 3.0 Hz, 1 H, H-4), 5.01 (dd, J = 10.9, 3.2 Hz, 1 H, H-3), 4.39 (t, J = 6.2 Hz, 1 H, H-5), 4.19-4.11 (m, 1 H, H-2), 4.11-3.96 (m, 2 H, H-6), 1.83, 1.93, 2.01, 2.12 (4 x s, 4 x 3 H, 4 x Ac); **^{13}C NMR** (DMSO- d_6 , 101 MHz) δ 170.56, 170.41, 170.36, 169.95 (4 x CO), 167.95 (CNH), 82.21 (C-1), 74.36 (C-5), 70.89 (C-3), 61.82 (C-6), 61.80 (C-4), 47.46 (C-2), 23.06 (NHAc), 20.89, 20.92, 21.05 (3 x OAc); **MS** (ESI $^+$) m/z = 406.43 [M+H]; **HRMS** (ESI $^+$) found 406.1278 [M+H], $\text{C}_{15}\text{H}_{24}\text{N}_3\text{O}_8\text{S}$ requires 406.1279; $[\alpha]_D^{20}$ = +4.3 (c = 1, DMSO); **IR** (ATR) 3219.73, 3187.60 (NH), 1749.11 (CO).

2-N-acetamido-3,4,6-tri-O-acetyl-1-thio-(S-cyanomethyl)-2-deoxy- β -D-galactopyranoside **2.45**



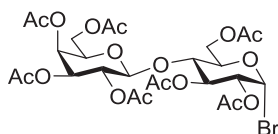
To a solution of **2.44** (2.269 g, 5.6 mmol, 1 eq.) in H₂O:acetone (1:1, 120 mL), were added Na₂S₂O₅ (2.34 g, 12.3 mmol, 2.2 eq.), K₂CO₃ (1.01 g, 7.28 mmol, 1.3 eq.) and chloroacetonitrile (6.73 mL, 106.4 mmol, 19 eq.). The mixture was stirred for 16 h at r.t., then poured into cold water (200 mL) and extracted with DCM (3 x 100 mL), the organic layer backwashed with brine (100 mL), dried over MgSO₄ and evaporated, to afford the title compound **2.45** as a white solid (1.2129 g, 3 mmol, 54%). **TLC** R_f = 0.55 (EtOAc, 100%); **¹H NMR** (CDCl₃, 400 MHz) δ 6.07 (d, *J* = 9.3 Hz, 1 H, NH), 5.42 (d, *J* = 2.8 Hz, 1 H, H-4), 5.19 (dd, *J* = 10.8, 3.3 Hz, 1 H, H-3), 4.83 (d, *J* = 10.3 Hz, 1 H, H-1), 4.32 (dd, *J* = 20.0, 10.4 Hz, 1 H, H-2), 4.23-4.10 (m, 2 H, H-6), 4.03 (t, *J* = 6.2 Hz, 1 H, H-5), 3.68 (d, *J* = 17.0 Hz, 1 H, CH₂aCN), 3.35 (d, *J* = 17.0 Hz, 1 H, CH₂bCN), 2.16, 2.05, 2.01, 1.98 (4 x s, 4 x 3 H, 4 x Ac); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.91, 170.76, 170.62, 170.28 (4 x CO), 116.43 (CN), 83.27 (C-1), 74.99 (C-5), 70.85 (C-3), 66.85 (C-4), 61.74 (C-6), 49.09 (C-2), 23.29 (NHAc), 20.74, 20.55, 20.12 (3 x OAc), 14.55 (CH₂CN); **MS** (ESI⁺) *m/z* = 404.41 [M+H]; **HRMS** (ESI⁺) found 402.1088 [M+H], C₁₆H₂₂N₂O₈S requires 402.1086; [α]_D²⁰ = -70.2 (*c* = 1, CDCl₃); **IR** (ATR) 2241.54 (CN), 1743.29 (CO).

**1,2,3,6-O-tetra-O-acetyl-4-O-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl]-
 β -D-glucopyranoside **2.46****



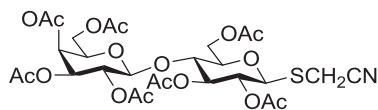
To β -lactose (10 g, 29.2 mmol, 1 eq.) were added Ac_2O (155 mL, 1636 mmol, 56 eq.) and NaOAc (10 g, 122.6 mmol, 4.2 eq.). The reaction was stirred at 140 °C for 5 h, recovered with H_2O (200 mL) and extracted with DCM (3 x 200 mL). The organic layer was washed with NaHCO_3 satd. solution (200 mL) and brine (200 mL), then dried over MgSO_4 , evaporated and co-evaporated with toluene, to afford the title compound **2.46** as a light yellow foam (19.9 g, 29.2 mmol, quant.). **TLC** R_f = 0.2 (PE:EtOAc, 1:1); **$^1\text{H NMR}$** (CDCl_3 , 400 MHz) δ 6.23 (d, J = 3.7 Hz, 0.3 H, H-1eq), 5.65 (d, J = 8.3 Hz, 0.7 H, H-1ax), 5.48-5.41 (m, 0.3 H, H-3 α), 5.34 (t, J = 3.2 Hz, 1 H, H-4'), 5.23 (t, J = 9.2 Hz, 0.7 H, H-3 β), 5.14-4.89 (m, 3 H, H-4, H-2, H-2'), 4.51-4.38 (m, 2 H, H-5, H-5'), 4.17-4.02 (m, 2 H-6a, H-6'a), 3.92-3.70 (m, 2 H, H-6b, H-6'b), 2.21, 2.14, 2.10, 2.08, 2.05, 2.03, 2.01, 1.95 (8 x s, 8 x 3 H, 8 x OAc); **$^{13}\text{C NMR}$** (CDCl_3 , 101 MHz) δ 170.76, 170.72, 170.54, 170.38, 170.27, 170.11, 169.84, 168.52 (8 x CO), 101.37 (C-1'), 87.22 (C-1), 74.33 (C-4), 72.35 (C-5), 71.52 (C-3'), 70.34 (C-5'), 70.11 (C-2), 69.34 (C-3), 69.12 (C-2'), 67.32 (C-4'), 60.92 (C-6'), 60.34 (C-6), 20.97, 20.88, 20.76, 20.71, 20.69, 20.65, 20.53, 20.12 (8 x Ac); **MS** (ESI^+) m/z = 679.7 [$\text{M}+\text{H}$]; $[\alpha]_D^{20}$ = -32.9 (c = 1, CHCl_3);^[22] **IR** (ATR) 1744.92 (CO).

2,3,6-O-tri-O-acetyl-4-O-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl]-1-bromo- α -D-glucopyranoside **2.47^[23]**



To **2.46** (14.15 g, 20.85 mmol, 1 eq.) in an ice bath, were added HBr (33% in AcOH, 23 mL, 129.27 mmol, 6.2 eq.) and Ac₂O (4.2 mL, 43.785 mmol, 2.1 eq.). The reaction was stirred for 3 h, recovered with DCM (200 mL) and poured into H₂O (200 mL). The mixture was stirred with NaHCO₃ satd. solution (300 mL), the organic layer separated and washed with NaHCO₃ satd. solution (3 x 200 mL) and brine (200 mL). The organic layer was dried over MgSO₄ and evaporated to afford the title compound **2.47** as a white foam (13.27 g, 19 mmol, 91%). **TLC** R_f = 0.3 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 6.50 (d, *J* = 4.0 Hz, 1 H, H-1), 5.52 (t, *J* = 9.6 Hz, 1 H, H-3), 5.33 (d, *J* = 2.9 Hz, 1 H, H-4'), 5.10 (dd, *J* = 10.4, 7.9 Hz, 1 H, H-2'), 4.93 (dd, *J* = 10.4, 3.4 Hz, 1 H, H-3'), 4.73 (dd, *J* = 10.0, 4.0 Hz, 1 H, H-2), 4.49 (d, *J* = 7.9 Hz, 2 H, H-1', H-6a), 4.29-3.97 (m, 4 H, H-6b, H-6', H-5), 3.93-3.76 (m, 2 H, H-5', H-4), 2.13, 2.11, 2.07, 2.04, 2.04, 2.03, 1.94 (7 x s, 7 x 3 H, 7 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.37, 170.20, 170.17, 170.10, 170.00, 169.26, 168.98 (7 x CO), 100.83 (C-1'), 86.43 (C-1), 74.98 (C-4), 72.99 (C-5), 71.00 (C-3'), 70.85 (C-5'), 70.78 (C-2), 69.59 (C-3), 69.01 (C-2'), 66.62 (C-4'), 61.06 (C-6'), 60.90 (C-6), 20.84, 20.81, 20.74, 20.69, 20.68, 20.858, 20.53 (7 x Ac); **MS** (ESI⁺) *m/z* = 700.46 [M+H]⁺; [α]_D²⁰ = -33.1 (c = 1, CHCl₃);^[24] **IR** (ATR) 1743.78 (CO).

2,3,6-O-tri-O-acetyl-4-O-[2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl]-1-thio-(S-cyanomethyl)-β-D-glucopyranoside 2.48

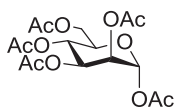


2.47 (14.58 g, 20.85 mmol, 1 eq.) was mixed with thiourea (3.5 g, 45.87 mmol, 2.2 eq.) and dissolved in acetone (50 mL). The reaction was stirred for 3 h at 80 °C, then evaporated.

The solid was recovered with acetone (60 mL), and Na₂S₂O₅ (10.3 g, 54.21 mmol, 2.6 eq.), K₂CO₃ (4 g, 29.19 mmol, 1.4 eq.) and chloroacetonitrile (20 mL, 316.92 mmol, 15.2 eq.) were added. The mixture was stirred at r.t. for 3 h, evaporated and purified by flash column chromatography on silica (EtOAc:PE, 1:1) to afford the title compound **2.48** as a white foam (8.7621 g, 12.6 mmol, 63%). **TLC** R_f = 0.25 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 5.34 (d, *J* = 2.6 Hz, 1 H, H-4'), 5.24 (t, *J* = 9.2 Hz, 1 H, H-3), 5.10 (dd, *J* = 10.4, 7.9 Hz, 1 H, H-2'), 5.02-4.88 (m, 2 H, H-2, H-3'), 4.68 (d, *J* = 10.0 Hz, 1 H, H-1), 4.54 (dd, *J* = 12.1, 1.9 Hz, 1 H, H-6a), 4.48 (d, *J* = 7.9 Hz, 1 H, H-1'), 4.19-4.02 (m, 3 H, H-6b, H-6'), 3.87 (t, *J* = 6.8 Hz, 1 H, H-5'), 3.82 (t, *J* = 9.5 Hz, 1 H, H-4), 3.68 (ddd, *J* = 9.9, 5.0, 1.9 Hz, 1 H, H-5), 3.58 (d, *J* = 17.0 Hz, 1 H, CH₂aCN), 3.29 (d, *J* = 17.0 Hz, 1 H, CH₂bCN), 2.14, 2.12, 2.09, 2.06, 2.06, 2.04, 1.95 (7 x s, 7 x 3 H, 7 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.38, 170.16, 170.10, 169.81, 169.57, 169.32, 169.03 (7 x CO), 115.74 (CN), 101.06 (C-1'), 81.48 (C-1), 75.79 (C-5), 73.25 (C-4), 70.94 (C-3), 70.78 (C-5'), 69.89, 69.87 (C-3', C-2), 69.04 (C-2'), 66.61 (C-4'), 61.63 (C-6'), 60.83 (C-6), 20.87, 20.81, 20.76, 20.71, 20.68, 20.65, 20.54 (7 x OAc), 14.48

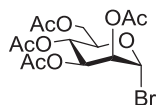
(CH₂CN); **MS** (ESI⁺) *m/z* = 692.66 [M+H]; **HRMS** (ESI⁺) found 692.1887 [M+H], C₂₈H₃₈NO₁₇S requires 692.1855; $[\alpha]_D^{20} = -34.1$ (*c* = 1, CHCl₃); **IR** (ATR) 22141.58 (CN), 1743.41 (CO).

1,2,3,4,6-penta-O-acetyl- α -D-mannopyranoside 2.49^[25]



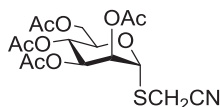
To acetic anhydride (200 mL) was added D-mannopyranoside (5 g, 27.75 mmol, 1 eq.) under nitrogen atmosphere and with rapid stirring, followed by pyridine (100 mL) and DMAP (678 mg, 5.55 mmol, 0.2 eq.). The mixture was stirred for 1 h at r.t.. After this time the reaction was slowly quenched with MeOH (200 mL), poured into ice (200 g) and H₂O (50 mL) and recovered with EtOAc (200 mL). The organic layer was separated and washed with HCl 1M (250 mL), NaHCO₃ satd. solution (250 mL), brine (250 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum to afford the title compound **2.49** as a yellow oil (10.8 g, 27.66 mmol, quant.). **TLC** *R_f* = 0.4 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 6.08 (d, *J* = 1.9 Hz, 1 H, H-1), 5.34 (dd, *J* = 3.6, 2.6 Hz, 2 H, H-3, H-4), 5.28-5.24 (m, 1 H, H-2), 4.28 (dd, *J* = 12.4, 4.9 Hz, 1 H, H-6a), 4.11 (dd, *J* = 4.8, 2.3 Hz, 1 H, H-6b), 4.04 (ddd, *J* = 8.1, 4.6, 2.4 Hz, 1 H, H-5), 2.17, 2.17, 2.09, 2.05, 2.00 (5 x s, 5 x 3 H, 5 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.99, 170.31, 169.84, 169.52, 169.03 (5 x CO), 83.84 (C-1), 70.07 (C-5), 69.33 (C-2), 69.06 (C-3), 65.58 (C-4), 61.87 (C-6), 20.91, 20.80, 20.75, 20.64, 20.53 (5 x OAc); **MS** (ESI⁺) *m/z* = 391.35 [M+H]; $[\alpha]_D^{20} = -75.6$ (*c* = 1, CHCl₃);^[26] **IR** (ATR) 1744.77 (CO).

2,3,4,6-tetra-O-acetyl-1-bromo- α -D-mannopyranoside **2.50**^[27]



To **2.49** (10.8 g, 27.75 mmol, 1 eq.) was added HBr (33% in AcOH, 40 mL) and acetic anhydride (10 mL). The mixture was stirred for 1 h at r.t.. After this time the reaction was poured into ice (200 g) and H₂O (50 mL), recovered with EtOAc (200 mL) and separated. The organic layer was washed with H₂O (250 mL), NaHCO₃ satd. solution (250 mL), brine (250 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum to afford the title compound **2.50** as a yellow oil (5.8 g, 14.87 mmol, 52%). **TLC** R_f = 0.45 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 6.29 (d, J = 1.1 Hz, 1 H, H-1), 5.71 (dd, J = 10.2, 3.4 Hz, 1 H, H-3), 5.44 (dd, J = 3.4, 1.6 Hz, 1 H, H-2), 5.37 (t, J = 10.2 Hz, 1 H, H-4), 4.33 (dd, J = 12.5, 4.9 Hz, 1 H, H-6a), 4.26-4.18 (m, 1 H, H-5), 4.12 (dd, J = 4.7, 2.4 Hz, 1 H, H-6b), 2.17, 2.10, 2.07, 2.01 (4 x s, 4 x 3 H, 4 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.41, 170.23, 169.71, 169.13 (4 x CO), 83.91 (C-1), 70.57 (C-5), 69.12 (C-2), 69.00 (C-3), 66.44 (C-4), 62.54 (C-6), 20.88, 20.73, 20.58, 20.50 (4 x OAc); cchsqc=171.3 Hz; **MS** (ESI⁺) m/z = 412.21 [M+H]; $[\alpha]_D^{20}$ = +111.2 (c = 1, CHCl₃);^[28] **IR** (ATR) 1746.43 (CO).

2,3,4,6-tetra-O-acetyl-1-thio-(S-cyanomethyl)- α -D-mannopyranoside **2.51**^[4]



To **2.50** (5.8 g, 14.1 mmol, 1 eq.) in acetone (50 mL) was added thiourea (2.36 g, 31 mmol, 2.2 eq.) and the mixture refluxed for 2 h at 80 °C. After this time the reaction was cooled to r.t. and the solvent removed under vacuum.

The solid was dissolved in acetone:H₂O (1:1, 100 mL) and K₂CO₃ (2.9 g, 21.15 mmol, 1.5 eq.), Na₂S₂O₅ (7.2 g, 38.07 mmol, 2.7 eq.) and chloroacetonitrile (4.5 mL, 70.5 mmol, 5 eq.) were added. The mixture was stirred for 12 h at r.t.. After this time the acetone was evaporated, and the aqueous layer extracted with DCM (3 x 50 mL). The organic layer was washed with brine (125 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum and purified by flash column chromatography on silica (EtOAc:PE, 1:1). The product was re-crystallised from hot MeOH, washed with cold Et₂O and dried to afford the title compound **2.51** as a white solid (2.33 g, 5.77 mmol, 27%). **TLC** R_f = 0.75 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 5.46 (d, *J* = 1.4 Hz, 1 H, H-1), 5.38-5.32 (m, 2 H, H-2, H-4), 5.19 (dd, *J* = 10.0, 3.5 Hz, 1 H, H-3), 4.39-4.28 (m, 2 H, H-6a, H-5), 4.16 (dd, *J* = 11.8, 1.6 Hz, 1 H, H-6b), 3.47 (d, *J* = 17.2 Hz, 1 H, CH₂aCN), 3.31 (d, *J* = 17.2 Hz, 1 H, CH₂bCN), 2.18, 2.11, 2.06, 2.00 (4 x s, 4 x 3 H, 4 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.55, 169.83, 169.74, 169.59 (4 x CO), 115.66 (CN), 82.44 (C-1), 69.95 (C-5), 69.57 (C-2), 69.26 (C-3), 65.80 (C-4), 62.01 (C-6), 20.82, 20.76, 20.68, 20.60 (4 x OAc), 15.79 (CH₂CN);

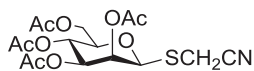
cchsqc=171.3 Hz; **MS** (ESI⁺) m/z = 404.41 [M+H]; $[\alpha]_D^{20} = +122.5$ (c = 1, CHCl₃); **IR** (ATR) 1744.75 (CO).

2,3,4,6-tetra-O-acetyl-1-thio-(S-acetyl)- β -D-mannopyranoside 2.52^[29]



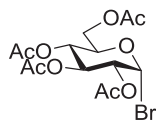
To a solution of thioacetic acid (557.5 μ L, 7.8 mmol, 2 eq.) in DMF (dry, 6 mL) at 0 °C under nitrogen was added potassium tert-butoxide (525 mg, 4.68 mmol, 1.2 eq.) portion-wise, and the mixture stirred at r.t. for 30 min. After this time **2.50** (1.6 g, 3.9 mmol, 1 eq.) in THF (dry, 2 mL) was added dropwise and the reaction stirred at r.t. for 16 h. The reaction was then diluted with EtOAc (50 mL) and washed with H₂O (200 mL) and brine (200 mL). The organic layer was dried over MgSO₄ and the solvent removed under vacuum. The crude was purified by flash column chromatography on silica (EtOAc:PE, 1:4 to 1:2) to afford the title compound **2.52** as a yellow solid (703.8 mg, 1.73 mmol, 44%). **TLC** R_f = 0.5 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 5.51 (ddd, *J* = 5.4, 4.2, 2.2 Hz, 2 H, H-1, H-2), 5.26 (t, *J* = 10.0 Hz, 1 H, H-4), 5.17 (dd, *J* = 10.1, 3.4 Hz, 1 H, H-3), 4.27 (dd, *J* = 8.5, 3.8 Hz, 1 H, H-6a), 4.12 (dd, *J* = 10.8, 3.5 Hz, 1 H, H-6b), 3.85 (ddd, *J* = 9.8, 5.2, 2.2 Hz, 1 H, H-5), 2.38, 2.11, 2.09, 2.06, 1.98 (5 x s, 5 x 3 H, 5 x Ac); **¹³C NMR** (CDCl₃, 101 MHz) δ 191.61 (CO-SAc), 169.96, 169.93, 169.79, 169.64 (4 x CO-OAc), 79.23 (C-1), 76.74 (C-5), 71.67 (C-3), 70.57 (C-2), 65.21 (C-4), 62.53, 62.29 (C-6), 30.56 (SAc), 20.69, 20.62, 20.54, 20.46 (4 x OAc); **MS** (ESI⁺) m/z = 407.41 [M+H]; $[\alpha]_D^{20} = -80.7$ (c = 1, CHCl₃); **IR** (ATR) 1746.12 (CO).

2,3,4,6-tetra-O-acetyl-1-thio-(S-cyanomethyl)- β -D-mannopyranoside **2.53**^[30]



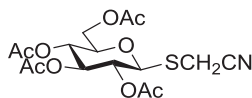
To **2.52** (703.8 mg, 1.73 mmol, 1 eq.) in MeOH (20 mL) at -40 °C was added NaOMe (25% in MeOH, 277 μ L, 1.21 mmol, 0.7 eq.) in MeOH (10 mL). The reaction was stirred at -40 °C for 15 min before being quenched with DowexH⁺, filtered and evaporated. The crude was recovered with DCM (20 mL), and DIPEA (904 μ L, 5.19 mmol, 3 eq.) and chloroacetonitrile (3.28 mL, 51.9 mmol, 30 eq.) were added. The reaction was stirred for 4 h at r.t.. After this time the solvent was evaporated, and the crude purified by flash column chromatography on silica (PE:EtOAc, 1:1 to 0:1) to afford the title compound **2.53** as light yellow oil (346 mg, 0.86 mmol, 50%). **TLC** R_f = 0.45 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 500 MHz) δ 5.54 (d, J = 3.2 Hz, 1 H, H-2), 5.28 (t, J = 10.1 Hz, 1 H, H-4), 5.15 (dd, J = 10.1, 3.4 Hz, 1 H, H-3), 5.08 (br, 1 H, H-1), 4.29 (dd, J = 12.4, 5.8 Hz, 1 H, H-6a), 4.17 (dd, J = 12.3, 1.9 Hz, 1 H, H-6b), 3.88-3.71 (m, 1 H, H-5), 3.63 (d, J = 17.1 Hz, 1 H, CH₂aCN), 3.40 (d, J = 17.1 Hz, 1 H, CH₂bCN), 2.19, 2.10, 2.06, 1.99 (4 x s, 4 x 3 H, 4 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.63, 170.00, 169.90, 169.61 (4 x CO), 115.87 (CN), 80.64 (C-1), 76.89 (C-5), 71.57 (C-3), 69.52 (C-2), 65.49 (C-4), 62.38 (C-6), 20.74, 20.66, 20.56, 20.51 (4 x OAc), 15.47 (CH₂CN); cchsqc=156.3 Hz; **MS** (ESI⁺) m/z = 404.41 [M+H]; $[\alpha]_D^{20}$ = -81.8 (c = 1, CHCl₃); **IR** (ATR) 1744.75 (CO).

2,3,4,6-tetra-O-acetyl-1-bromo- α -D-glucopyranoside **2.54**^[31]



To 1,2,3,4,6-penta-O-acetyl- β -D-glucopyranoside (3 g, 7.69 mmol, 1 eq.) was added HBr (33% in AcOH, 12 mL) and acetic anhydride (3 mL). The mixture was stirred for 1 h at r.t.. After this time the reaction was poured into ice (20 g) and H₂O (50 mL), recovered with EtOAc (150 mL) and separated. The organic layer was washed with H₂O (125 mL), NaHCO₃ satd. solution (125 mL), brine (125 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum to afford the title compound **2.54** as a colourless oil (3.5 g, 7.6 mmol, quant.). **TLC** R_f = 0.9 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 6.53 (d, *J* = 4.0 Hz, 1 H, H-1), 5.46 (t, *J* = 9.7 Hz, 1 H, H-3), 5.08 (t, *J* = 9.8 Hz, 1 H, H-4), 4.76 (dd, *J* = 10.0, 4.0 Hz, 1 H, H-2), 4.31-4.17 (m, 2 H, H-5, H-6a), 4.06-3.98 (m, 1 H, H-6b), 2.02, 2.01, 1.97, 1.95 (4 x s, 4 x 3 H, 4 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.11, 170.07, 169.83, 169.39 (4 x CO), 81.72 (C-1), 77.94 (C-5), 72.16 (C-3), 70.48 (C-2), 68.52 (C-4), 60.12 (C-6), 20.78, 20.59, 20.43, 20.12 (4 x OAc); **MS** (ESI⁺) *m/z* = 412.21 [M+H]; [α]_D²⁰ = +80.4 (*c* = 1, CHCl₃);^[32] **IR** (ATR) 1746.05 (CO).

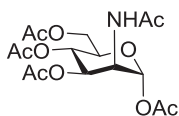
2,3,4,6-tetra-O-acetyl-1-thio-(S-cyanomethyl)- α -D-glucopyranoside **2.55**^[4]



To **2.54** (3.5 g, 7.69 mmol, 1 eq.) in acetone (20 mL) was added thiourea (1.17 g, 15.38 mmol, 2 eq.) and the mixture refluxed for 2 h at 80 °C. After this time the reaction was cooled to r.t. and the solvent removed under vacuum.

The solid was recovered with acetone:H₂O (1:1, 40 mL) and K₂CO₃ (1.6 g, 11.53 mmol, 1.5 eq.), Na₂S₂O₅ (3.95 g, 20.76 mmol, 2.7 eq.) and chloroacetonitrile (2.4 mL, 38.45 mmol, 5 eq.) were added. The mixture was stirred for 12 h at r.t.. After this time the acetone was evaporated, and the aqueous layer extracted with DCM (2 x 100 mL). The organic layer was washed with brine (125 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum and the product recrystallised from hot Et₂O to afford the title compound **2.55** as a white solid (2 g, 4.96 mmol, 65%). **TLC** R_f = 0.85 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 5.29 (t, *J* = 9.4 Hz, 1 H, H-3), 5.13 (t, *J* = 9.8 Hz, 1 H, H-4), 5.06 (t, *J* = 9.7 Hz, 1 H, H-2), 4.74 (d, *J* = 10.0 Hz, 1 H, H-1), 4.23 (ddd, *J* = 14.7, 12.5, 3.5 Hz, 2 H, H-6), 3.80 (ddd, *J* = 10.1, 4.8, 2.3 Hz, 1 H, H-5), 3.65 (d, *J* = 17.0 Hz, 1 H, CH₂aCN), 3.35 (d, *J* = 17.0 Hz, 1 H, CH₂bCN), 2.11, 2.09, 2.05, 2.03 (4 x s, 4 x 3 H, 4 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.66, 170.03, 169.59, 169.42 (4 x CO), 115.74 (CN), 81.78 (C-1), 76.28 (C-5), 73.38 (C-3), 69.51 (C-2), 67.96 (C-4), 61.69 (C-6), 20.75, 20.66, 20.60, 20.58 (4 x OAc), 14.59 (CH₂CN); **MS** (ESI⁺) *m/z* = 404.41 [M+H]; [α]_D²⁰ = -61.7 (c = 1, CHCl₃);^[4] **IR** (ATR) 1743.44 (CO).

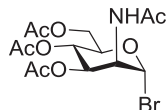
2-N-acetamido-1,3,4,6-tetra-O-acetyl-2-deoxy- α -D-mannopyranoside **2.56**^[33]



To acetic anhydride (30 mL) was added 2-acetamido-2-deoxy- β -D-mannopyranoside (1 g, 4.5 mmol, 1 eq.), followed by pyridine (20 mL) and DMAP (110 mg, 0.9 mmol, 0.2 eq.). The mixture was stirred for 2 h at r.t.. After this time the reaction was poured into ice (20 g) and H₂O (50 mL) and recovered with IPA:DCM (1:4, 50 mL). The organic layer was separated and washed with HCl 1M (250 mL), NaHCO₃ satd. solution (250 mL), brine (250 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum to afford the title compound **2.56** as a colourless oil (1.76 g, 4.5 mmol, quant). **TLC** R_f = 0.7 (EtOAc, 100%); **¹H NMR** (CDCl₃, 400 MHz) δ 5.92 (d, *J* = 9.1 Hz, 1 H, H-1), 5.85 (d, *J* = 1.7 Hz, 1 H, NH), 5.13 (t, *J* = 9.8 Hz, 1 H, H-4), 5.05 (dd, *J* = 9.9, 4.0 Hz, 1 H, H-3), 4.77 (ddd, *J* = 9.1, 4.0, 1.7 Hz, 1 H, H-2), 4.28 (dd, *J* = 12.5, 5.2 Hz, 1 H, H-6a), 4.08 (dd, *J* = 12.4, 2.4 Hz, 1 H, H-6b), 3.79 (ddd, *J* = 9.6, 5.2, 2.4 Hz, 1 H, H-5), 2.10, 2.09, 2.09, 2.06, 2.00 (5 x s, 5 x 3 H, 5 x Ac); **¹³C NMR** (CDCl₃, 101 MHz) δ 175.81 (CO-NHAc), 170.95, 170.14, 169.80, 168.38 (4 x CO-OAc), 90.69 (C-1), 73.50 (C-5), 71.33 (C-3), 65.31 (C-4), 61.91 (C-6), 49.60 (C-2), 23.34 (NHAc), 20.88, 20.78, 20.74, 20.68 (4 x OAc); **MS** (ESI⁺) *m/z* = 390.36 [M+H]; $[\alpha]_D^{20}$ = +14.9 (*c* = 1, CHCl₃); **IR** (ATR) 1744.58 (CO).

2-N-acetamido-3,4,6-tri-O-acetyl-1-bromo-2-deoxy- α -D-mannopyranoside

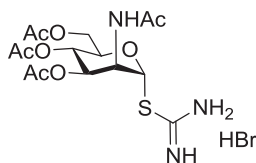
2.57



To **2.56** (1.76 g, 4.5 mmol, 1 eq.) was added HBr (33% in AcOH, 2 mL, 11.25 mmol, 2.5 eq.) and acetic anhydride (340 μ L, 3.6 mmol, 0.8 eq.). The mixture was stirred for 1.5 h at r.t.. After this time the reaction was evaporated and then co-evaporated with DCM to afford the title compound **2.57** as a yellow oil (1.8 g, 4.39 mmol, 98%). **TLC** R_f = 0.2 (EtOAc, 100%); **$^1\text{H NMR}$** (CDCl_3 , 400 MHz) δ 6.39 (s, 1 H, NH), 5.71 (dd, J = 10.2, 4.6 Hz, 1 H, H-3), 5.44-5.36 (m, 1 H, H-4), 4.83 (dd, J = 8.1, 4.7 Hz, 1 H, H-2), 4.42 (dd, J = 12.4, 5.7 Hz, 1 H, H-6a), 4.31-4.23 (m, 1 H, H-5), 4.11 (dd, J = 12.4, 2.3 Hz, 1 H, H-6b), 2.09, 2.09, 2.08, 2.03 (4 x s, 4 x 3 H, 4 x Ac); **$^{13}\text{C NMR}$** (CDCl_3 , 101 MHz) δ 176.72 (CO-NHAc), 170.15, 169.54, 169.31 (3 x CO-OAc), 83.53 (C-1), 73.49 (C-5), 67.75 (C-3), 61.79 (C-4), 55.17 (C-6), 50.54 (C-2), 22.33 (NHAc), 20.83, 20.79, 20.74 (3 x OAc); **MS** (ESI^+) m/z = 411.22 $[\text{M}+\text{H}]$; $[\alpha]_D^{20}$ = +17.3 (c = 1, CHCl_3); **IR** (ATR) 1744.97 (CO).

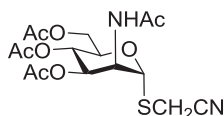
2-N-acetamido-3,4,6-tri-O-acetyl-2-deoxy- β -D-mannopyranosyl

thiouronium bromide **2.58**



To **2.57** (1.8 g, 4.5 mmol, 1 eq.) in acetone (5 mL) was added thiourea (754 mg, 9.9 mmol, 2.2 eq.) and the mixture refluxed for 2 h at 80 °C. After this time the reaction was cooled to r.t., filtered, washed with cold acetone and the solvent removed under vacuum to afford the title compound **2.58** as a pink solid (1.095 g, 2.7 mmol, 60%). **TLC** R_f = 0.0 (EtOAc, 100%); **$^1\text{H NMR}$** (DMSO- d_6 , 400 MHz) δ 8.99 (d, J = 5.2 Hz, 2 H, NH_2Br), 8.69-8.60 (m, 1 H, NHAc), 8.13 (dd, J = 7.6, 6.7 Hz, 2 H, NH_2), 6.15 (d, J = 2.0 Hz, 1 H, H-1), 5.33 (t, J = 9.1 Hz, 1 H, H-4), 5.01 (dd, J = 9.3, 4.6 Hz, 1 H, H-3), 4.56 (ddd, J = 8.9, 4.6, 2.1 Hz, 1 H, H-2), 4.28-4.21 (m, 2 H, H-5, H-6a), 4.14-4.02 (m, 1 H, H-6b), 1.96, 1.93, 1.89 1.87 (4 x Ac); **$^{13}\text{C NMR}$** (DMSO- d_6 , 101 MHz) δ 170.62, 170.24, 170.04, 169.98 (4 x CO), 142.46 (CNH_2Br), 127.76 (CNH_2), 83.01 (C-1), 69.13 (C-3), 67.07 (C-4), 63.86 (C-6), 50.80 (C-5), 49.28 (C-2), 21.38 (NHAc), 20.95, 20.90, 20.85 (3 x OAc); **MS** (ESI^+) m/z = 406.43 [$\text{M}+\text{H}$]; **HRMS** (ESI^+) found 406.1277 [$\text{M}+\text{H}$], $\text{C}_{15}\text{H}_{24}\text{N}_3\text{O}_8\text{S}$ requires 406.1279; $[\alpha]_D^{20}$ = +19.5 (c = 1, H_2O); **IR** (ATR) 1744.06 (CO).

2-N-acetamido-3,4,6-tri-O-acetyl-1-thio-(S-cyanomethyl)-2-deoxy-β-D-mannopyranoside 2.59



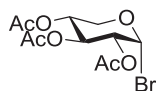
To **2.58** (613.7 mg, 1.5 mmol, 1 eq.) in acetone:H₂O (1:1, 14 mL) were added K₂CO₃ (311 mg, 2.25 mmol, 1.5 eq.), Na₂S₂O₅ (770 mg, 4.05 mmol, 2.7 eq.) and chloroacetonitrile (475 μL, 7.5 mmol, 5 eq.). The mixture was stirred for 1 h at r.t.. After this time the acetone was evaporated, and the aqueous layer extracted with DCM:IPA (3:1, 50 mL). The organic layer was washed with brine (125 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum to afford the title compound **2.59** as a white solid (414.2 mg, 1.03 mmol, 69%). **TLC** R_f = 0.6 (EtOAc, 100%); **¹H NMR** (CDCl₃, 400 MHz) δ 6.26 (d, *J* = 8.5 Hz, 1 H, NH), 5.35 (d, *J* = 1.3 Hz, 1 H, H-1), 5.19-5.03 (m, 2 H, H-3, H-4), 4.64 (d, *J* = 8.6 Hz, 1 H, H-2), 4.39-4.19 (m, 2 H, H-5, H-6a), 4.05 (dd, *J* = 13.6, 3.8 Hz, 1 H, H-6b), 3.40 (d, *J* = 16.8 Hz, 1 H, CH₂aCN), 3.26 (d, *J* = 17.2 Hz, 1 H, CH₂bCN), 2.05, 2.01, 2.00, 1.94 (4 x s, 4 x 3 H, 4 x Ac); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.65, 170.29, 169.88, 169.86 (4 x CO), 116.02 (CN), 83.42 (C-1), 69.57 (C-5), 69.44 (C-4), 66.03 (C-3), 62.28 (C-6), 50.19 (C-2), 25.32 (NHAc), 23.09, 20.77, 20.67 (3 x OAc), 15.74 (CH₂CN); cchsqc = 170.6 Hz; **MS** (ESI⁺) *m/z* = 403.43 [M+H]; **HRMS** (ESI⁺) found 403.1174 [M+H], C₁₆H₂₃N₂O₈S requires 403.1170; [α]_D²⁰ = +22.5 (c = 1, CHCl₃); **IR** (ATR) 1746.76 (CO).

1,2,3,4-tetra-O-acetyl- β -D-xylopyranoside **2.60**^[34]



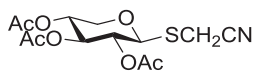
To acetic anhydride (35 mL) was added β -D-xylopyranoside (1 g, 6.66 mmol, 1 eq.) followed by pyridine (20 mL) and DMAP (163 mg, 1.33 mmol, 0.2 eq.). The mixture was stirred for 1 h at r.t.. After this time the reaction was poured into ice (20 g) and H₂O (50 mL) and recovered with EtOAc (150 mL). The organic layer was separated and washed with HCl 1M (250 mL), NaHCO₃ satd. solution (250 mL), brine (250 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum to afford the title compound **2.60** as a yellow oil (2.282 g, 6.6 mmol, quant). **TLC** R_f = 0.8 (PE:EtOAc, 1:2); **¹H NMR** (CDCl₃, 400 MHz) δ 6.24 (d, J = 3.6 Hz, 1 H, H-1), 5.46 (t, J = 9.8 Hz, 1 H, H-3), 5.10-4.94 (m, 2 H, H-2, H-4), 3.92 (dd, J = 11.2, 5.9 Hz, 1H, H-5a), 3.70 (t, J = 11.0 Hz, 1 H, H-5b), 2.16, 2.04, 2.04, 2.01 (4 x s, 4 x 3 H, 4 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.16, 169.80, 169.76, 169.05 (4 x CO), 89.25 (C-1), 69.35, 69.32 (C-3, C-4), 68.68 (C-2), 60.65 (C-5), 20.89, 20.76, 20.69, 20.52 (4 x OAc); **MS** (ESI⁺) m/z = 319.29 [M+H]; $[\alpha]_D^{20}$ = -58.52 (c = 1, CHCl₃);^[35] **IR** (ATR) 1744.61 (CO).

2,3,4-tri-O-acetyl-1-bromo- α -D-xylopyranoside 2.61^[36]



To **2.60** (2.28 g, 6.66 mmol, 1 eq.) was added HBr (33% in AcOH, 2.2 mL) and acetic anhydride (8.8 mL). The mixture was stirred for 1 h at r.t.. After this time the reaction was poured into ice (20 g) and H₂O (50 mL), recovered with EtOAc (150 mL) and separated. The organic layer was washed with H₂O (250 mL), NaHCO₃ satd. solution (250 mL), brine (250 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum to afford the title compound **2.61** as a yellow oil (2.236 g, 6.59 mmol, 99%). **TLC** R_f = 0.75 (PE:EtOAc, 1:2); **¹H NMR** (CDCl₃, 200 MHz) δ 6.54 (d, J = 4.0 Hz, 1 H, H-1), 5.51 (t, J = 9.8 Hz, 1 H, H-3), 5.13-4.88 (m, 1 H, H-4), 4.73 (dd, J = 9.9, 4.0 Hz, 1 H, H-2), 4.16-3.94 (m, 1 H, H-5a), 3.83 (dd, J = 13.5, 8.7 Hz, 1 H, H-5b), 2.05, 2.01 (3 x s, 3 x 3 H, 3 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.72, 169.84, 169.72 (3 x CO), 81.97 (C-1), 70.44 (C-3), 69.82 (C-4), 66.24 (C-2), 65.92 (C-5), 20.34, 20.09, 20.00 (3 x OAc); **MS** (ESI⁺) m/z = 340.15 [M+H]; $[\alpha]_D^{20}$ = -58.92 (c = 1, CHCl₃); **IR** (ATR) 1744.18 (CO).

2,3,4-tri-O-acetyl-1-thio-(S-cyanomethyl)- β -D-xylopyranoside **2.62**

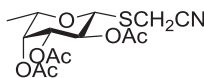


To **2.61** (2.236 g, 6.59 mmol, 1 eq.) in acetone (25 mL) was added thiourea (1.104 g, 14.5 mmol, 2.2 eq.) and the mixture refluxed for 2.5 h at 80 °C. After this time the reaction was cooled to r.t. and the solvent removed under vacuum.

The solid was recovered with acetone:H₂O (1:1, 20 mL) and K₂CO₃ (945 mg, 6.84 mmol, 1.5 eq.), Na₂S₂O₅ (2.34 g, 12.3 mmol, 2.7 eq.) and chloroacetonitrile (1.5 mL, 22.8 mmol, 5 eq.) were added. The mixture was stirred for 10 h at r.t.. After this time the acetone was evaporated, and the aqueous layer extracted with EtOAc (150 mL). The organic layer was washed with brine (125 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum to afford the title compound **2.62** as a white solid (1.3687 g, 4.1 mmol, 63%). **TLC** R_f = 0.95 (PE:EtOAc, 1:2); **¹H NMR** (CDCl₃, 400 MHz) δ 5.21 (t, J = 8.0 Hz, 1 H, H-3), 4.95 (dd, J = 10.3, 5.7 Hz, 2 H, H-4, H-2), 4.80 (d, J = 8.1 Hz, 1 H, H-1), 4.27 (dd, J = 11.9, 4.9 Hz, 1 H, H-5a), 3.59 (d, J = 17.0 Hz, 1 H, CH₂aCN), 3.48 (dd, J = 11.8, 8.5 Hz, 1 H, H-5b), 3.30 (d, J = 17.0 Hz, 1 H, CH₂bCN), 2.09, 2.07, 2.06 (3 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 169.79, 169.66, 169.57 (3 x CO), 115.95 (CN), 82.36 (C-1), 71.18 (C-3), 69.16 (C-4), 68.18 (C-2), 65.16 (C-5), 20.75, 20.69, 20.67 (3 x OAc), 14.97 (CH₂CN); **MS** (ESI⁺) m/z = 332.35 [M+H]; **HRMS** (ESI⁺) found 332.0797 [M+H], C₁₃H₁₈NO₇S requires 332.0798; $[\alpha]_D^{20}$ = -62.5 (c = 1, CHCl₃); **IR** (ATR) 1745.82 (CO).

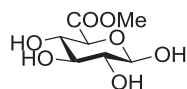
2,3,4-tri-O-acetyl-6-deoxy-1-thio-(S-cyanomethyl)- β -L-galactopyranoside

2.63



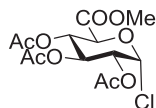
To **2.17** (1.2765 g, 3.6 mmol, 1 eq.) in ACN (6 mL) was added thiourea (289 mg, 3.8 mmol, 1.05 eq.) and the mixture refluxed for 1 h at 80 °C. After this time the reaction was cooled to r.t. and NEt₃ (1.25 mL, 9 mmol, 2.5 eq.) and chloroacetonitrile (342 μ L, 5.4 mmol, 1.5 eq.) were added. The mixture was stirred at r.t. for other 1 h. The solvent was evaporated, the reaction mixture recovered with EtOAc (150 mL) and the organic layer washed with H₂O (125 mL), HCl 1M (125 mL), brine (125 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum and the crude filtered over silica plug with EtOAc to afford the title compound **2.63** as a white solid (1.1683 g, 3.39 mmol, 94%). **TLC** R_f = 0.9 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 5.29-5.22 (m, 1 H, H-4), 5.16 (t, J = 9.9 Hz, 1 H, H-2), 5.03 (dd, J = 10.0, 3.3 Hz, 1 H, H-3), 4.61 (d, J = 9.8 Hz, 1 H, H-1), 3.90-3.76 (m, 1 H, H-5), 3.59 (d, J = 17.0 Hz, 1 H, CH₂aCN), 3.28 (d, J = 17.0 Hz, 1 H, CH₂bCN), 2.13, 2.02, 1.93 (3 x OAc), 1.18 (d, J = 6.4 Hz, 3 H, H-6); **¹³C NMR** (CDCl₃, 101 MHz) δ 170.02, 170.00, 169.94 (3 x CO), 115.99 (CN), 81.97 (C-1), 73.79 (C-5), 71.91 (C-3), 70.13 (C-4), 66.80 (C-2), 20.77, 20.70, 20.63 (3 x OAc), 16.22 (C-6), 14.39 (CH₂CN); **MS** (ESI⁺) m/z = 346.38 [M+H]; $[\alpha]_D^{20}$ = +22.2 (c = 1, CHCl₃); **IR** (ATR) 1741.89 (CO).

β -D-glucuronic acid methyl ester **2.64**^[37]



To β -D-glucuronic acid (0.5 g, 2.57 mmol, 1 eq.) in MeOH (40 mL) was added DowexH⁺ (1.3 g). The suspension was stirred for 3 h at r.t.. After this time the reaction was filtered, the resin washed with MeOH, and the solution was concentrated under vacuum. Finally, the crude was filtered over silica plug with EtOAc:MeOH (9:1) to afford the title compound **2.64** as a colourless oil (573.7 mg, 2.42 mmol, 94%). **TLC** R_f = 0.9 (MeOH:EtOAc, 1:1); **¹H NMR** (MeOD, 400 MHz) δ 5.56-5.44 (br, 1 H, H-1), 5.07-4.84 (m, 2 H, H-2, H-3), 3.75-3.59 (m, 2 H, H-5, H-4), 3.68 (s, 3 H, Me); **¹³C NMR** (MeOD, 101 MHz) δ 170.16 (COOMe), 100.17 (C-1), 79.63 (C-5), 74.87 (C-3), 73.17 (C-2), 69.99 (C-4), 52.28 (OMe); **MS** (ESI⁻) m/z = 221.18 [M-H]; $[\alpha]_D^{20}$ = -94.6 (c = 1, D₂O); **IR** (ATR) 1763.23 (CO).

2,3,4-tri-O-acetyl-1-chloro- α -D-glucuronic acid methyl ester **2.65**^[38]

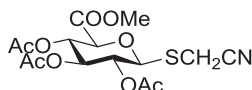


2.64 (573.7 mg, 2.76 mmol, 1 eq.) was dissolved in AcOH (14 mL) and AcCl (20 mL) was added. The mixture was stirred for 22 h at r.t.. After this time the reaction was evaporated and then co-evaporated with chloroform to afford the title compound **2.65** as a yellow oil (720 mg, 2.14 mmol, 77%). **TLC** R_f = 0.85 (EtOAc, 128

100%); **¹H NMR** (CDCl₃, 400 MHz) δ 5.26-5.20 (m, 2 H, H-3, H-4), 5.02-4.96 (m, 2 H, H-1, H-2), 3.75 (d, *J* = 3.2 Hz, 1 H, H-5), 3.39 (s, 3 H, Me), 2.25, 2.22, 2.12 (3 x s, 3 x 3 H, 3 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 169.75, 169.43, 169.32 (3 x s, 3 x 3 H, 3 x OAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 169.75, 169.43, 169.32 (3 x CO-OAc), 166.37 (COOMe), 82.09 (C-1), 76.77 (C-5), 72.93 (C-3), 69.46 (C-2), 68.89 (C-4), 53.00 (OMe), 20.72, 20.66, 20.51 (3 x OAc); **MS** (ESI⁻) *m/z* = 335.73 [M-H]; [α]_D²⁰ = +34.5 (*c* = 1, CHCl₃);^[39] **IR** (ATR) 1742.96 (CO).

2,3,4-tri-O-acetyl-1-thio-(S-cyanomethyl)-β-D-glucuronic acid methyl ester

2.66

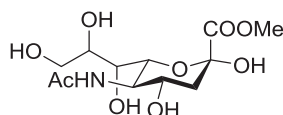


2.65 (300 g, 0.75 mmol, 1 eq.) was recovered with acetone (5 mL) and thiourea (115 mg, 1.51 mmol, 2 eq.) was added. The mixture was refluxed for 2 h at 80 °C. After this time the reaction was cooled to r.t. and the solvent removed under vacuum.

The solid was dissolved in acetone:H₂O (1:1, 5 mL) and K₂CO₃ (156 mg, 1.13 mmol, 1.5 eq.), Na₂S₂O₅ (388 mg, 2.04 mmol, 2.7 eq.) and chloroacetonitrile (240 μL, 3.775 mmol, 5 eq.) were added. The mixture was stirred for 18 h at r.t.. After this time the acetone was evaporated, and the aqueous layer extracted with EtOAc (150 mL). The organic layer was washed with brine (125 mL) and dried over MgSO₄. Finally, the solution was concentrated under vacuum and purified by flash column chromatography on silica (PE:EtOAc, 2:1 to 3:2) to afford the title

compound **2.66** as a white solid (66 mg, 0.17 mmol, 23%). **TLC** R_f = 0.5 (PE:EtOAc, 1:1); **$^1\text{H NMR}$** (CDCl_3 , 400 MHz) δ 5.30 (t, J = 9.3 Hz, 1 H, H-3), 5.21 (t, J = 9.7 Hz, 1 H, H-4), 5.03 (t, J = 9.5 Hz, 1 H, H-2), 4.75 (d, J = 10.0 Hz, 1 H, H-1), 4.09 (d, J = 9.8 Hz, 1 H, H-5), 3.73 (s, 3 H, OMe), 3.64 (d, J = 17.0 Hz, 1 H, CH_2aCN), 3.33 (d, J = 17.0 Hz, 1 H, CH_2bCN), 2.05, 2.00, 2.00 (3 x s, 3 x 3 H, 3 x OAc); **$^{13}\text{C NMR}$** (CDCl_3 , 101 MHz) δ 169.86, 169.49, 169.36 (3 x CO-OAc), 166.54 (COOMe), 115.67 (CN), 81.97 (C-1), 76.10 (C-5), 72.53 (C-3), 69.22 (C-2), 69.08 (C-4), 53.08 (OMe), 20.58, 20.51, 20.48 (3 x OAc), 14.75 (CH_2CN); **MS** (ESI^+) m/z = 390.39 [$\text{M}+\text{H}$]; **HRMS** (ESI^+) found 390.0855 [$\text{M}+\text{H}$], $\text{C}_{15}\text{H}_{20}\text{NO}_9\text{S}$ requires 390.0853; $[\alpha]_D^{20}$ = -37.5 (c = 1, CHCl_3); **IR** (ATR) 1747.17 (CO).

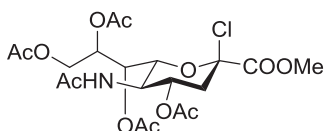
N-acetyl- α -D-neuraminic acid methyl ester **2.67**^[40]



To N-acetyl- α -D-neuraminic acid (0.5 g, 1.62 mmol, 1 eq.) in MeOH (40 mL) was added DowexH⁺ (0.7 g). The mixture was stirred for 3 h at r.t.. After this time the reaction was filtered, the resin washed with MeOH, and the solution was concentrated under vacuum. Finally, the crude was filtered over silica plug with EtOAc:MeOH (9:1) to afford the title compound **2.67** as a white solid (449.5 mg, 1.39 mmol, 86%). **TLC** R_f = 0.6 (MeOH:EtOAc, 1:1); **$^1\text{H NMR}$** (MeOD, 400 MHz) δ 4.14-3.96 (m, 2 H, H-4, H-6), 3.88-3.76 (m, 5 H, H-5, H-9a, Me), 3.71 (ddd, J = 8.7, 5.6, 2.8 Hz, 1 H, H-8), 3.63 (dd, J = 11.2, 5.6 Hz, 1 H, H-9b), 3.52-3.45 (m, 1

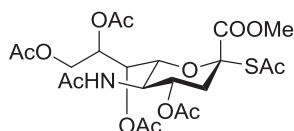
H, H-7), 2.22 (dd, $J = 12.9, 4.9$ Hz, 1 H, H-3a), 2.02 (s, 3 H, Ac), 1.96-1.83 (m, 1 H, H-3b); ^{13}C NMR (MeOD, 101 MHz) δ 170.15 (CO-NHAc), 166.34 (COOMe), 82.96 (C-2), 76.35 (C-6), 70.12 (C-4), 67.31 (C-8), 66.69 (C-7), 62.82 (C-9), 53.01 (Me), 48.07 (C-5), 36.11 (C-3), 23.49 (NHAc); **MS** (ESI $^-$) $m/z = 322.29$ [M-H]; $[\alpha]_{\text{D}}^{20} = -107.2$ ($c = 1$, MeOH); $^{[41]}$ **IR** (ATR) 1748.91 (CO).

N-acetyl-4,7,8,9-tetra-O-acetyl-2-chloro- β -D-neuraminic acid methyl ester
2.68



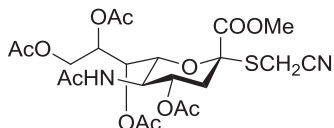
To **2.67** (449.5 mg, 1.39 mmol, 1 eq.) in AcOH (11 mL) was added AcCl (16 mL). The mixture was stirred for 1 h at r.t.. After this time the reaction was evaporated and co-evaporated with chloroform to afford the title compound **2.68** as a yellow oil (968 mg, 1.3 mmol, quant.). **TLC** $R_f = 0.2$ (EtOAc, 100%); ^1H NMR (CDCl_3 , 400 MHz) δ 5.47 (dd, $J = 7.2, 2.4$ Hz, 1 H, H-4), 5.44-5.34 (m, 1 H, H-7), 5.23-5.13 (m, 1 H, H-8), 4.42 (dd, $J = 12.5, 2.7$ Hz, 1 H, H-9a), 4.35 (dd, $J = 10.8, 2.4$ Hz, 1 H, H-6), 4.28 (br, 1 H, H-5), 4.06 (dd, $J = 12.6, 5.7$ Hz, 1 H, H-9b), 3.88 (s, 3 H, Me), 2.79 (dd, $J = 13.8, 4.8$ Hz, 1 H, H-3a), 2.28 (dd, $J = 13.9, 11.2$ Hz, 1 H, H-3b), 2.12, 2.08, 2.06, 2.05 (4 x s, 4 x 3 H, 4 x OAc), 1.91 (s, 3 H, NHAc); ^{13}C NMR (CDCl_3 , 101 MHz) δ 170.98, 170.72, 170.64, 170.25, 169.81 (5 x CO-Ac), 166.91 (COOMe), 82.32 (C-2), 74.58 (C-6), 69.47 (C-4), 67.05 (C-8), 66.35 (C-7), 62.94 (C-9), 53.39 (Me), 48.92 (C-5), 36.22 (C-3), 23.44 (NHAc), 21.92, 21.13, 20.92, 20.27 (4 x OAc); **MS** (ESI $^-$) $m/z = 508.88$ [M-H]; $[\alpha]_{\text{D}}^{20} = +103.6$ ($c = 1$, CHCl_3); **IR** (ATR) 1742.78 (CO).

N-acetyl-4,7,8,9-tetra-O-acetyl-2-thioacetyl- β -D-neuraminic acid methyl ester **2.69^[42]**



To **2.68** (968 mg, 1.39 mmol, 1 eq.) in DCM (10 mL) was added potassium thioacetate (875 mg, 7.66 mmol, 5.5 eq.) and the mixture was stirred at r.t. for 16 h. After this time the reaction was washed with H₂O (125 mL), brine (125 mL), dried over MgSO₄ and evaporated. The crude was purified by flash column chromatography on silica (EtOAc, 100%) to afford the title compound **2.69** as a yellow oil (448.4 mg, 0.82 mmol, 59%). **TLC** R_f = 0.3 (EtOAc, 100%); **¹H NMR** (CDCl₃, 400 MHz) δ 5.65 (d, J = 10.2 Hz, 1 H, NH), 5.33 (dd, J = 6.0, 2.4 Hz, 1 H, H-7), 5.19 (td, J = 6.2, 2.4 Hz, 1 H, H-8), 4.87 (td, J = 11.6, 4.6 Hz, 1 H, H-4), 4.65 (dd, J = 10.8, 2.4 Hz, 1 H, H-6), 4.38 (dd, J = 12.4, 2.5 Hz, 1 H, H-9a), 4.14-4.04 (m, 2 H, H-5, H-9b), 3.75 (s, 3 H, Me), 2.58 (dd, J = 12.9, 4.6 Hz, 1 H, H-3a), 2.24 (s, 3 H, SAc), 2.10, 2.08 (2 x s, 2 x 3 H, 2 x OAc), 2.06-2.03 (m, 1 H, H-3b), 2.00, 1.99 (2 x s, 2 x 3 H, 2 x OAc), 1.84 (s, 3 H, NHAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 171.33, 170.82, 170.75, 170.31, 170.22, 169.61 (6 x CO-Ac), 167.23 (COOMe), 82.67 (C-2), 74.45 (C-6), 69.94 (C-4), 67.31 (C-8), 66.45 (C-7), 62.54 (C-9), 53.99 (Me), 48.83 (C-5), 36.52 (C-3), 35.02 (SAc), 23.84 (NHAc), 21.17, 21.01, 20.61, 20.51 (4 x OAc); **MS** (ESI⁺) m/z = 550.55 [M+H]; $[\alpha]_D^{20}$ = +109.2 (c = 1, CHCl₃); **IR** (ATR) 1744.63 (CO).

N-acetyl-4,7,8,9-tetra-O-acetyl-2-thio-(S-cyanomethyl)- β -D-neuraminic acid methyl ester 2.70

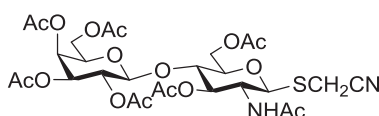


To **2.69** (448.4 mg, 0.816 mmol, 1 eq.) in MeOH (10 mL) at $-40\text{ }^{\circ}\text{C}$ was added NaOMe (25% in MeOH, 131 μL in 4 mL MeOH, 0.57 mmol, 0.7 eq.) and stirred for 15 min. The mixture was quenched with DowexH⁺, filtered and evaporated. To the crude were added 10 mL of DCM, DIPEA (426 μL , 2.45 mmol, 3 eq.) and chloroacetonitrile (1.6 mL, 24.48 mmol, 30 eq.) and stirred for 16 h at r.t.. After this time the reaction was concentrated under vacuum and purified by flash column chromatography on silica (EtOAc, 100%) to afford the title compound **2.70** as a white solid (321 mg, 0.59 mmol, 72%). **TLC** R_f = 0.2 (EtOAc, 100%); **¹H NMR** (CDCl₃, 400 MHz) δ 5.66 (d, J = 10.0 Hz, 1 H, NH), 5.36 (ddd, J = 8.6, 5.6, 2.7 Hz, 1 H, H-8), 5.24 (dd, J = 9.2, 2.1 Hz, 1 H, H-7), 4.81 (td, J = 11.3, 4.6 Hz, 1 H, H-4), 4.16 (dd, J = 12.4, 2.7 Hz, 1 H, H-9a), 4.01-3.93 (m, 2 H, H-9b, H-5), 3.87-3.82 (m, 1 H, H-6), 3.81 (s, 3 H, Me), 3.66 (d, J = 17.8 Hz, 1 H, CH₂aCN), 3.33 (d, J = 17.8 Hz, 1 H, CH₂bCN), 2.68 (dd, J = 12.5, 4.6 Hz, 1 H, H-3a), 2.12, 1.97, 1.96, 1.95 (4 x s, 4 x 3 H, 4 x OAc), 1.89-1.81 (m, 1 H, H-3b), 1.79 (s, 3 H, NHAc); **¹³C NMR** (CDCl₃, 101 MHz) δ 171.29, 170.65, 170.46, 170.37, 170.02 (5 x CO-Ac), 167.79 (COOMe), 116.95 (CN), 82.07 (C-2), 74.19 (C-6), 69.17 (C-4), 67.54 (C-8), 66.85 (C-7), 62.39 (C-9), 53.56 (Me), 48.82 (C-5), 36.96 (C-3), 23.05, 21.20, 21.03, 20.84, 20.79 (5 x Ac), 13.76 (CH₂CN); hsqmbc ³J_C-H (**H-3b**-

COOMe)= 7.8 Hz; **MS** (ESI⁺) *m/z* = 547.55 [M+H]; **HRMS** (ESI⁺) found 547.1591 [M+H], C₂₂H₃₁N₂O₁₂S requires 547.1592; [α]_D²⁰ = +131.2 (c = 1, CHCl₃); **IR** (ATR) 2252.12 (CN), 1737.55 (CO).

2-N-acetamido-3,6-di-O-acetyl-4-O-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside

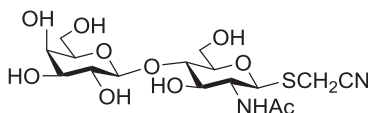
2.71



To **2.31** (664 mg, 0.96 mmol, 1 eq.) were added pyridine (4 mL), Ac₂O (10 mL) and DMAP (22 mg, 0.19 mmol, 0.2 eq.). The reaction was stirred at r.t. for 2.5 h, then poured into ice (20 g) and H₂O (50 mL), recovered with EtOAc (150 mL) and separated. The organic layer was washed with HCl 1 M (150 mL), NaHCO₃ satd. solution (150 mL), brine (150 mL) and dried over MgSO₄. Finally, the solution was evaporated and co-evaporated with toluene to afford the title compound **2.71** as a light yellow solid (210 mg, 0.3 mmol, 32%). **TLC** *R*_f = 0.85 (EtOH:CHCl₃, 1:10); **¹H NMR** (CDCl₃, 400 MHz) δ 6.18 (d, *J* = 9.6 Hz, 1 H, NH), 5.37-5.26 (m, 1 H, H-4'), 5.20-4.97 (m, 2 H, H-3, H-2'), 4.91 (dd, *J* = 10.4, 3.4 Hz, 1 H, H-3'), 4.65-4.58 (m, 1 H, H-1), 4.55-4.38 (m, 2 H, H-1', H-6a), 4.14 (d, *J* = 9.8 Hz, 1 H, H-2), 4.03 (dd, *J* = 12.1, 6.8 Hz, 3 H, H-6b, H-6'), 3.86 (dt, *J* = 13.9, 7.0 Hz, 1 H, H-5'), 3.77 (t, *J* = 9.3 Hz, 1 H, H-4), 3.67-3.56 (m, 2 H, CH₂aCN, H-5), 3.22 (d, *J* = 17.0 Hz, 1 H, CH₂bCN), 2.09, 2.06, 2.02, 2.00, 1.99, 1.91, 1.90 (7 x s, 7 x 3 H, 7 x Ac); **¹³C**

NMR (CDCl₃, 101 MHz) δ 170.94, 170.60, 170.57, 170.44, 170.37, 170.14, 169.22 (7 x CO), 129.02 (CN), 101.10 (C-1'), 82.77 (C-1), 76.97 (C-5), 75.85 (C-4), 73.69 (C-3), 70.89 (C-3'), 70.72 (C-5'), 69.09 (C-2'), 66.66 (C-4'), 61.76 (C-6), 60.84 (C-6'), 52.43 (C-2), 23.04 (NHAc), 20.85, 20.80, 20.66, 20.63, 20.61, 20.50 (6 x OAc), 14.52 (CH₂CN); **MS** (ESI⁺) m/z = 691.68 [M+H]⁺; $[\alpha]_D^{20}$ = -30.7 (c = 1, CHCl₃); **IR** (ATR) 2249.46 (CN), 1742.58 (CO).

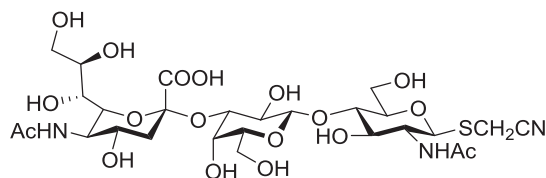
2-N-acetamido-4-O-[β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside **2.72**



To **2.31** (1.1 g, 1.6 mmol, 1 eq.) in MeOH (20 mL), was added triethylamine (1.1 mL, 8 mmol, 5 eq.). The reaction was stirred at r.t. for 48 h, then evaporated to afford the title compound **2.72** as a light yellow solid (660 mg, 1.5 mmol, 94%). **TLC** R_f = 0.0 (EtOH:CHCl₃, 1:10); **¹H NMR** (DMSO-d₆, 400 MHz) δ 8.01 (t, J = 8.8 Hz, 1 H, NH), 4.54 (dt, J = 30.9, 15.5 Hz, 2 H, H-1, H-1'), 4.36 (d, J = 10.3 Hz, 1 H, OH), 4.20 (ddd, J = 20.5, 10.9, 6.8 Hz, 2 H, H-5, H-5'), 3.98-3.05 (m, 17 H, H-2, H-2', H-3, H-3', H-4, H-4', H-6, H-6', 5 x OH, CH₂CN), 1.79 (s, 3 H, NHAc); **¹³C NMR** (DMSO-d₆, 101 MHz) δ 169.74 (CO), 104.51 (CN), 83.27, 79.77, 77.85, 76.30, 76.16, 73.60, 73.53, 72.24, 71.03, 70.81, 68.66, 60.99 (C-1, C-1', C-2, C-2', C-3, C-3', C-4, C-4', C-5, C-5', C-6, C-6'), 22.98 (NHAc), 12.39 (CH₂CN); **MS** (ESI⁺) m/z = 439.46 [M+H]⁺; $[\alpha]_D^{20}$ = -28.2 (c = 1, D₂O); **IR** (ATR) 3359.65 (OH), 1582.67 (CO).

2-N-acetamido-4-O-[3-O-(N-Acetyl- α -D-neuraminic acid)- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside

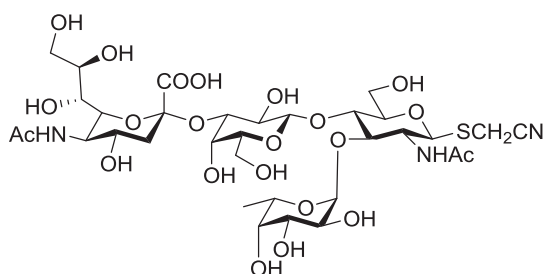
2.73



To **2.72** (80 mg, 0.186 mmol, 1 eq.) were added CTP (13 mM, 0.24 mmol, 1.3 eq.), sialic acid (13 mM, 0.24 mmol, 1.3 eq.), MgCl_2 (15 mM, 0.276 mmol, 1.5 eq.), Na_2CO_3 (50 mM, 0.92 mmol, 5 eq.) and the pH was adjusted to 8.5. To the reaction mixture were added NmCSS (4 mg/mL, 10 μL , final concentration 10 $\mu\text{g/mL}$), PmST (4 mg/mL, 30 μL , final concentration 30 $\mu\text{g/mL}$) and the reaction was stirred at 37 $^\circ\text{C}$ for 4 h. The reaction mixture was lyophilised, recovered with MeOH, filtered, evaporated and purified by flash column chromatography on silica (WIPE, 1:2:2) to afford the title compound **2.73** as a white solid (100 mg, 0.14 mmol, 76%). **TLC** R_f = 0.15 (WIPE, 1:2:2); **^1H NMR** (MeOD, 500 MHz) δ 4.70 (d, J = 10.4 Hz, 1 H, H-1), 4.44 (t, J = 24.3 Hz, 1 H, H-1'), 4.06 (dd, J = 9.7, 2.9 Hz, 1 H, H-4), 4.03-3.80 (m, 8 H, NH, H-9a, H-6a, H-3, H-2, H-6', CH_2aCN), 3.80-3.56 (m, 12 H, H-5', H-9b, H-6b, H-6''', H-3', H-4', CH_2bCN , H-8, H-2', H-3''', H-5''', NH), 3.56-3.46 (m, 2 H, H-5, H-7), 2.85 (dd, J = 12.2, 3.7 Hz, 1 H, H-3'''a), 2.03, 1.99 (2 x s, 2 x 3 H, 2 x NHAc), 1.75 (t, J = 11.6 Hz, 1 H, H-3'''b); **^{13}C NMR** (MeOD, 126 MHz) δ 174.11 (CO-NHAc), 173.73 (COOH), 172.32 (CO-NHAc), 117.23 (CN), 103.55 (C-1'), 99.74 (C-2'''), 83.49 (C-1), 79.46 (C-7), 79.25 (C-2'), 76.09 (C-4), 75.57 (C-3'), 73.70 (C-4'), 73.49 (C-5'), 71.61 (C-2), 69.39 (C-4'''), 68.57 (C-5), 67.58 (C-8), 62.87 (C-9), 61.29 (C-6), 60.59 (C-6'), 53.80 (C-3),

52.57 (C-6'''), 48.54 (C-5'''), 40.52 (C-3'''), 21.47, 21.39 (2 x NHAc), 13.66 (CH₂CN); **MS** (ESI⁻) *m/z* = 728.7 [M-H]; [α]_D²⁰ = -19.5 (c = 1, D₂O); **IR** (ATR) 3357.49 (OH), 1583.42 (CO).

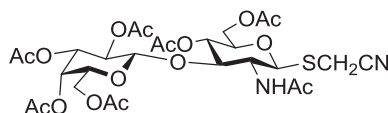
2-N-acetamido-3-O-[fucopyranosyl]-4-O-[3-O-(N-Acetyl- α -D-neuraminic acid)- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside 2.74



To **2.73** (60 mg, 0.081 mmol, 1 eq.) were added GDP-Fuc (0.219 mmol, 2.7 eq.), MgCl₂ (15 mM, 0.12 mmol, 1.5 eq.) and the pH was adjusted to 8. To the reaction mixture was added FucT (0.55 mg/mL, 1.65 mL, final concentration 150 μ g/mL) and the reaction stirred at 37 °C for 3 h. The reaction mixture was lyophilised, recovered with MeOH, filtered and evaporated to title compound **2.74** as a white solid (100.4 mg, 0.08 mmol, quant). **TLC** *R_f* = 0.1 (WIPE, 1:2:2); **¹H NMR** (MeOD, 500 MHz) δ 5.06 (d, *J* = 3.9 Hz, 1 H, H-1''), 4.83-4.79 (m, 1 H, H-5''), 4.72 (d, *J* = 10.1 Hz, 1 H, H-1), 4.50 (d, *J* = 7.8 Hz, 1 H, H-1'), 4.14-4.07 (m, 1 H, H-2), 4.03 (dd, *J* = 9.7, 3.1 Hz, 1 H, H-6'a), 4.00-3.81 (m, 9 H, H-6'b, H-9a, CH₂aCN, H-3, H-4, H-3', H-4', H-3'', NH), 3.80-3.69 (m, 5 H, CH₂bCN, H-6, H-4''', H-4''), 3.68-3.59 (m, 5 H, H-9b, H-2'', H-6''', H-8, NH), 3.59-3.43 (m, 5 H, H-5, H-5', H-7, H-

2', H-5'''), 2.87 (dd, $J = 12.4, 4.1$ Hz, 1 H, H-3''')a), 2.02, 1.97 (2 x s, 2 x 3 H, 2 x Ac), 1.73 (t, $J = 11.9$ Hz, 1 H, H-3''')b), 1.17 (d, $J = 6.6$ Hz, 3 H, H-6''); **^{13}C NMR** (MeOD, 126 MHz) δ 174.08 (CO-NHAc), 173.50 (COOH), 172.38 (CO-NHAc), 117.22 (CN), 102.57 (C-1'), 99.49 (C-1''), 98.87 (C-2'''), 83.21 (C-1), 80.39 (C-7), 76.40, 76.06, 75.30, 73.73, 73.55, 72.27, 71.66, 69.55, 69.44, 68.64, 68.35, 67.86 (C-2', C-4, C-2'', C-4', C-3', C-5', C-3'', C-4'', C-4''', C-5, C-7, C-8, C-5''), 67.35 (C-5'), 66.36 (C-9), 63.04 (C-6), 61.66 (C-6'), 59.93 (C-2), 54.31 (C-3), 52.54 (C-6'''), 48.47 (C-5'''), 40.88 (C-3'''), 21.57, 21.26 (2 x Ac), 15.18 (C-6''), 13.44 (CH₂CN); **MS** (ESI⁻) $m/z = 874.84$ [M-H]; $[\alpha]_{\text{D}}^{20} = -17.7$ ($c = 1$, D₂O); **IR** (ATR) 3359.53 (OH), 2263.41 (CN), 1619.07 (CO).

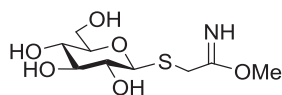
2-N-acetamido-4,6-di-O-acetyl-3-O-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl]-1-thio-(S-cyanomethyl)-2-deoxy- β -D-glucopyranoside
2.75



To **2.26** (855 mg, 0.96 mmol, 1 eq.) in DCM (dry, 10 mL) and pyridine (dry, 312 μL , 3.86 mmol, 4 eq.) was added AcCl (172 μL , 2.412 mmol, 2.5 eq.). After 1 h MeOH was added to quench the reaction and stirred for a few min at r.t.. The crude was evaporated, recovered with EtOAc (100 mL) and washed with HCl 1M (100 mL), NaHCO₃ (100 mL), brine (100 mL) and dried over MgSO₄. Evaporation of the solvent afforded the title compound **2.75** as a white amorphous solid (590

mg, 0.85 mmol, 89%). **TLC** R_f = 0.7 (MeOH:EtOAc, 0.5:10); **$^1\text{H NMR}$** (CDCl_3 , 400 MHz) δ 5.99 (d, J = 8.7 Hz, 1 H, NH), 5.35 (d, J = 3.0 Hz, 1 H, H-4'), 5.05 (dd, J = 10.4, 7.8 Hz, 1 H, H-2'), 5.00-4.88 (m, 3 H, H-3', H-1, H-3), 4.59 (d, J = 7.7 Hz, 1 H, H-1'), 4.28-4.00 (m, 5 H, H-4, H-6, H-6'), 3.89 (t, J = 6.9 Hz, 1 H, H-5'), 3.83 (d, J = 10.0 Hz, 1 H, H-2), 3.80-3.72 (m, 1 H, H-5), 3.55 (d, J = 17.2 Hz, 1 H, CH_2aCN), 3.39 (d, J = 17.2 Hz, 1 H, CH_2bCN), 2.14, 2.08, 2.07, 2.06, 2.05, 2.04, 1.96 (7 x s, 7 x 3 H, 7 x Ac); **$^{13}\text{C NMR}$** (CDCl_3 , 101 MHz) δ 170.76, 170.68, 170.48, 170.24, 170.18, 169.36, 169.32 (7 x CO), 116.83 (CN), 100.99 (C-1'), 81.64 (C-1), 78.19 (C-4), 76.37 (C-5), 70.92 (C-3'), 70.59 (C-5'), 69.13 (C-2'), 68.80 (C-3), 66.84 (C-4'), 62.28 (C-6'), 61.05 (C-6), 53.44 (C-2), 23.43 (NHAc), 20.80, 20.77, 20.71, 20.63, 20.55, 20.52 (6 x OAc), 14.05 (CH_2CN); **MS** (ESI^+) m/z = 691.68 [$\text{M}+\text{H}$]; **HRMS** (ESI^+) found 691.2016 [$\text{M}+\text{H}$], $\text{C}_{28}\text{H}_{39}\text{N}_2\text{O}_{16}\text{S}$ requires 691.2015; $[\alpha]_D^{20}$ = -33.4 (c = 1, CHCl_3); **IR** (ATR) 1744.78 (CO).

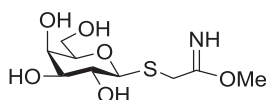
1-thio-(S-2-imido-2-methoxyethyl)- β -D-glucopyranoside **2.76**



To a solution of **2.55** (20 mg, 0.05 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 12 μL , 0.05 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.76**) in ratio 1:0.5 (determined by NMR, ^1H , MeOD), and used in the next step without further

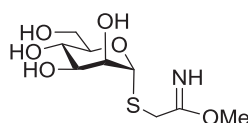
purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **^1H NMR** (D_2O , 500 MHz) δ 4.60 (d, J = 9.7 Hz, 1 H, H-1), 3.83 (dd, J = 12.4, 2.1 Hz, 1 H, H-6a), 3.64 (dd, J = 12.4, 5.9 Hz, 1 H, H-6b), 3.48-3.36 (m, 2 H, H-5, CH_2aCNH), 3.36-3.25 (m, 7 H, CH_2bCNH , H-2, H-4, H-3, OMe); **^{13}C NMR** (D_2O , 126 MHz) δ 85.04 (C-1), 80.24 (C-5), 77.95 (C-4), 72.54 (C-2), 69.68 (C-3), 60.99 (CH_2CNH), 60.95 (C-6), 48.83 (OMe).

1-thio-(S-2-imido-2-methoxyethyl)- β -D-galactopyranoside **2.77**



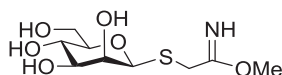
To a solution of **2.42** (20 mg, 0.05 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 12 μL , 0.05 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.77**) in ratio 1:0.7 (determined by NMR, ^1H , MeOD), and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **^1H NMR** (D_2O , 500 MHz) δ 4.56 (d, J = 9.1 Hz, 1 H, H-1), 3.90 (d, J = 2.9 Hz, 1 H, H-4), 3.75-3.45 (m, 7 H, H-5, H-6, CH_2CNH , H-3, H-2), 3.25 (s, 3 H, OMe); **^{13}C NMR** (D_2O , 126 MHz) δ 85.24 (C-1), 79.29 (C-5), 74.30 (C-3), 69.58 (C-2), 68.90 (C-4), 61.11 (CH_2CNH), 61.09 (C-6), 48.83 (OMe).

1-thio-(S-2-imido-2-methoxyethyl)- α -D-mannopyranoside **2.78**



To a solution of **2.51** (20 mg, 0.05 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 12 μ L, 0.05 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.78**) in ratio 1:0.5 (determined by NMR, ¹H, MeOD), and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 500 MHz) δ 5.51 (d, J = 0.9 Hz, 1 H, H-1), 4.10 (dd, J = 2.8, 1.5 Hz, 1 H, H-2), 3.99-3.93 (m, 2 H, H-4, H-6a), 3.92 (d, J = 2.3 Hz, 1 H, CH₂aCNH), 3.86 (d, J = 5.5 Hz, 1 H, CH₂bCNH), 3.85-3.80 (m, 1 H, H-6b), 3.80-3.76 (m, 2 H, H-5, H-3), 3.38 (s, 3 H, OMe); **¹³C NMR** (D₂O, 126 MHz) δ 175.06 (CNH), 85.29 (C-1), 73.69 (C-4), 71.65 (C-5), 71.20 (C-2), 67.06 (C-3), 60.73 (C-6), 60.64 (CH₂CNH), 48.83 (OMe).

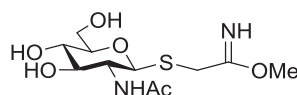
1-thio-(S-2-imido-2-methoxyethyl)- β -D-mannopyranoside **2.79**



To a solution of **2.53** (20 mg, 0.05 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 12 μ L, 0.05 mmol, 1 eq.). The reaction was stirred at r.t.

for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.79**) in ratio 1:0.6 (determined by NMR, ¹H, MeOD), and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 500 MHz) δ 4.76 (m, 1 H, H-1), 4.00-3.94 (m, 1 H, H-2), 3.88-3.78 (m, 1 H, H-6a), 3.69-3.47 (m, 5 H, CH₂CNH, H-6b, H-3, H-4), 3.36-3.20 (m, 4 H, H-5, OMe); **¹³C NMR** (D₂O, 126 MHz) δ 183.61 (CNH), 84.09 (C-1), 80.33 (C-5), 73.83 (C-3), 71.87 (C-2), 66.69 (C-4), 61.16 (CH₂CNH), 61.12 (C-6), 48.86 (OMe).

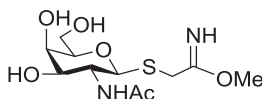
2-acetamido-1-thio-(S-2-imido-2-methoxyethyl)-2-deoxy-β-D-glucopyranoside **2.80**



To a solution of **2.3** (20 mg, 0.05 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 12 μL, 0.05 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.80**) in ratio 1:0.5 (determined by NMR, ¹H, MeOD), and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 500 MHz) δ 4.71 (d, *J* = 9.3 Hz, 1 H, H-1), 3.85 (dd, *J* = 12.4, 2.0 Hz, 1 H, H-6a), 3.72 (t, *J* = 10.1 Hz, 1 H, H-2), 3.67 (dd, *J* = 12.4, 5.4 Hz, 1 H, H-6b), 3.56-3.47 (m, 1 H, H-3),

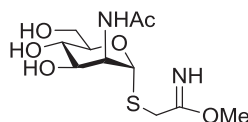
3.46-3.34 (m, 4 H, CH₂CNH, H-5, H-4), 3.26 (s, 3 H, OMe), 1.95 (s, 3 H, Ac); ¹³C NMR (D₂O, 126 MHz) δ 174.66 (CO), 118.43 (CN), 83.80 (C-1), 80.29 (C-5), 75.08 (C-3), 69.99 (C-4), 60.92, 60.89 (C-6, CH₂CNH), 54.34 (C-2), 48.83 (OMe), 22.07 (Ac).

2-acetamido-1-thio-(S-2-imido-2-methoxyethyl)-2-deoxy-β-D-galactopyranoside 2.81



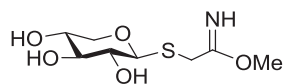
To a solution of **2.45** (20 mg, 0.05 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 12 μL, 0.05 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.81**) in ratio 1:0.6 (determined by NMR, ¹H, MeOD), and used in the next step without further purification. TLC R_f = 0.0 (DCM:MeOH, 10:1); ¹H NMR (D₂O, 500 MHz) δ 4.65 (d, J = 10.5 Hz, 1 H, H-1), 3.97-3.80 (m, 2 H, H-2, H-4), 3.76-3.56 (m, 6 H, H-3, CH₂CNH, H-6, H-5), 3.26 (s, 3 H, OMe), 1.95 (s, 3 H, Ac); ¹³C NMR (D₂O, 126 MHz) δ 174.86 (CNH), 84.15 (C-1), 79.31 (C-5), 71.88 (C-3), 68.04 (C-2), 61.15 (CH₂CNH), 61.12 (C-6), 50.94 (C-4), 48.83 (OMe), 22.14 (Ac).

2-acetamido-1-thio-(S-2-imido-2-methoxyethyl)-2-deoxy- α -D-mannopyranoside **2.82**



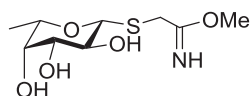
To a solution of **2.59** (20 mg, 0.05 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 12 μ L, 0.05 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.82**) in ratio 1:0.7 (determined by NMR, ¹H, MeOD), and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 500 MHz) δ 5.20 (m, 1 H, H-1), 4.33 (d, J = 4.6 Hz, 1 H, H-2), 4.00-3.82 (m, 2 H, H-3, H-5), 3.77 (dt, J = 20.3, 7.3 Hz, 4 H, CH₂CNH, H-6), 3.68-3.50 (m, 1 H, H-4), 3.27 (s, 3 H, OMe), 1.97 (s, 3 H, Ac); **¹³C NMR** (D₂O, 126 MHz) δ 174.46 (CNH), 83.68 (C-1), 72.95 (C-3), 69.76 (C-5), 66.98 (C-4), 60.33 (CH₂CNH), 60.20 (C-6), 53.43 (C-2), 48.85 (OMe), 21.86 (Ac).

1-thio-(S-2-imido-2-methoxyethyl)- β -D-xylopyranoside **2.83**



To a solution of **2.62** (20 mg, 0.06 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 14 μ L, 0.06 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.83**) in ratio 1:0.5 (determined by NMR, ¹H, MeOD), and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 500 MHz) δ 4.42 (d, *J* = 9.6 Hz, 1 H, H-1), 3.91 (ddd, *J* = 11.3, 5.4, 2.8 Hz, 1 H, H-5a), 3.57-3.52 (m, 1 H, H-4), 3.41-3.28 (m, 3 H, CH₂CNH, H-3), 3.28-3.26 (m, 4 H, OMe, H-5b), 3.24-3.21 (m, 1 H, H-2); **¹³C NMR** (D₂O, 126 MHz) δ 182.92 (CNH), 85.54 (C-1), 77.26 (C-3), 72.06 (C-2), 69.21 (CH₂CNH), 69.11 (C-4), 68.97 (C-5), 48.86 (OMe).

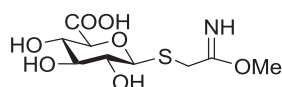
1-thio-(S-2-imido-2-methoxyethyl)- α -L-fucopyranoside **2.84**



To a solution of **2.63** (20 mg, 0.06 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 14 μ L, 0.06 mmol, 1 eq.). The reaction was stirred at r.t.

for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.84**) in ratio 1:0.5 (determined by NMR, ¹H, MeOD), and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 500 MHz) δ 4.55 (dd, *J* = 9.6, 0.9 Hz, 1 H, H-1), 3.77-3.66 (m, 3 H, H-5, H-4, CH₂aCNH), 3.64-3.41 (m, 3 H, CH₂bCNH, H-3, H-2), 3.26 (s, *J* = 1.0 Hz, 3 H, OMe), 1.18-1.11 (m, 3 H, H-6); **¹³C NMR** (D₂O, 126 MHz) δ 85.08 (C-1), 75.06 (C-5), 74.06 (C-3), 73.11 (C-4), 71.48 (C-2), 69.20 (CH₂CNH), 48.86 (OMe), 15.53 (C-6).

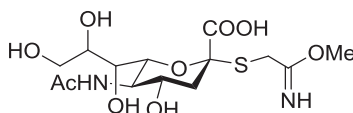
1-thio-(S-2-imido-2-methoxyethyl)-β-D-glucuronic acid **2.85**



To a solution of **2.66** (20 mg, 0.05 mmol, 1 eq.) in MeOH (dry, 0.5 mL) was added KOH (0.2 M, 0.5 mL). The reaction was stirred at r.t. for 10 min, then neutralised with DowexH⁺ and evaporated. The residue was recovered with MeOH (dry, 1 mL) and NaOMe (25% in MeOH, 18 μL, 0.05 mmol, 1 eq.) was added. The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.85**) in ratio 1:0.8 (determined by NMR, ¹H, MeOD), and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 400 MHz) δ 4.59 (d, *J* = 9.8 Hz, 1 H, H-1), 3.74 (d, *J* = 9.3 Hz, 1 H, H-5), 3.61-3.28

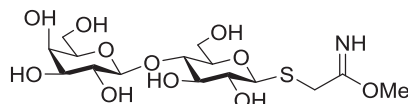
(m, 5 H, H-4, CH₂CNH, H-3, H-2); ¹³C NMR (D₂O, 101 MHz) δ 175.59 (CO), 84.91 (C-1), 80.05 (C-5), 77.03 (C-4), 72.19 (C-3), 71.80 (C-2), 71.71 (CH₂CNH).

2-thio-(S-2-imido-2-methoxyethyl)-N-acetyl-α-D-neuraminic acid **2.86**



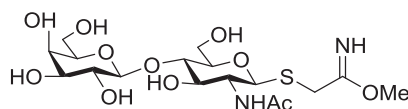
To a solution of **2.70** (20 mg, 0.04 mmol, 1 eq.) in MeOH (dry, 0.5 mL) was added KOH (0.2 M, 0.5 mL). The reaction was stirred at r.t. for 10 min, then neutralised with DowexH⁺ and evaporated. The residue was recovered with MeOH (dry, 1 mL) and NaOMe (25% in MeOH, 13 μL, 0.05 mmol, 1 eq.) was added. The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.86**) in ratio 1:0.6 (determined by NMR, ¹H, MeOD), and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); ¹H NMR (D₂O, 400 MHz) δ 4.42 (d, *J* = 8.6 Hz, 1 H, NH), 4.18 (d, *J* = 10.8 Hz, 1 H, H-6), 4.08-3.93 (m, 1 H, H-5), 3.93-3.47 (m, 7 H, H-7, H-9, H-4, CH₂CNH, H-8), 2.76 (dd, *J* = 12.5, 4.5 Hz, 1 H, H-3a), 1.98 (s, *J* = 4.4 Hz, 3 H, Ac), 1.76 (td, *J* = 11.9, 4.7 Hz, 1 H, H-3b); ¹³C NMR (D₂O, 101 MHz) δ 175.18 (CO), 85.84 (C-2), 75.34 (C-6), 74.95 (C-4), 71.80 (C-5), 68.52 (C-8), 68.26 (CH₂CNH), 62.67 (C-9), 51.67 (C-7), 40.52 (C-3), 22.04 (Ac).

4-O-[β -D-galactopyranosyl]-1-thio-(S-2-imido-2-methoxyethyl)- β -D-glucopyranoside **2.87**



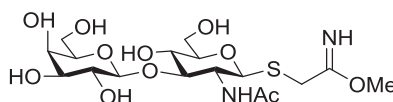
To a solution of **2.48** (100 mg, 0.14 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 33 μ L, 0.14 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.87**) in ratio 1:0.7 (determined by NMR, ¹H, MeOD), and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1). **¹H NMR** (D₂O, 500 MHz) δ 4.51 (d, J = 9.9 Hz, 1 H, H-1), 4.37 (d, J = 7.8 Hz, 1 H, H-1'), 3.87-3.82 (m, 2 H, H-6a, H-4'), 3.70 (ddd, J = 11.6, 8.8, 3.9 Hz, 4 H, H-6', H-6b, CH₂aCNH), 3.66-3.61 (m, 3 H, CH₂bCNH, H-3, H-5), 3.61-3.55 (m, 3 H, H-5', H-4, H-3'), 3.48-3.44 (m, 1 H, H-2'), 3.36-3.30 (m, 1 H, H-2), 3.26 (s, 3 H, OMe); **¹³C NMR** (D₂O, 126 MHz) δ 175.53 (CNH), 118.56 (CN), 102.81 (C-1'), 84.87 (C-1), 78.68 (C-3), 77.82 (C-3'), 77.75, 75.61, 75.32, 72.47 (C-4, C-4', C-5, C-5'), 71.83 (C-2'), 70.91 (C-2), 68.52 (CH₂CNH), 61.01 (C-6), 59.98 (C-6'), 48.82 (OMe).

2-N-acetamido-4-O-[β -D-galactopyranosyl]-1-thio-(S-2-imido-2-methoxyethyl)-2-deoxy- β -D-glucopyranoside **2.88**



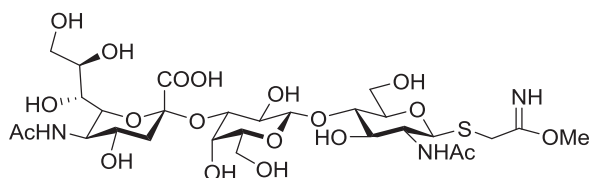
To a solution **2.71** (20 mg, 0.03 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 7 μ L, 0.03 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.88**) in ratio 1:0.7 (determined by NMR, ¹H, MeOD) , and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 400 MHz) δ 4.79 (d, *J* = 10.3 Hz, 1 H, H-1), 4.45 (d, *J* = 7.7 Hz, 1 H, H-1'), 4.05-3.59 (m, 13 H, H-6, H-2, H-3, H-4, H-5, H-4', H-5', H-6', CH₂CNH, H-3'), 3.52 (dd, *J* = 20.0, 12.1 Hz, 1 H, H-2'), 2.00 (s, 3 H, Ac); **¹³C NMR** (D₂O, 101 MHz) δ 174.69 (CO), 102.91 (C-1'), 83.77 (C-1), 78.94 (C-3), 78.07 (C-4), 75.42 (C-5), 73.45 (C3'), 72.59 (C-4'), 70.97 (C-5'), 68.62 (C-2'), 68.44 (CH₂CNH), 61.07 (C-6'), 60.11 (C-6), 53.91 (C-2), 22.12 (Ac).

2-N-acetamido-3-O-[β -D-galactopyranosyl]-1-thio-(S-2-imido-2-methoxyethyl)-2-deoxy- β -D-glucopyranoside **2.89**



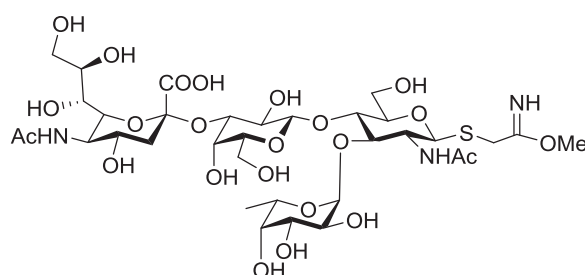
To a solution of **2.75** (20 mg, 0.03 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 7 μ L, 0.03 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.89**) in ratio 1:0.7 (determined by NMR, ¹H, MeOD) , and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 400 MHz) δ 4.72 (d, *J* = 10.4 Hz, 1 H, H-1), 4.34 (d, *J* = 7.7 Hz, 1 H, H-1'), 3.92-3.77 (m, 3 H, H-2, H-6a, H-4), 3.77-3.70 (m, 1 H, H-3), 3.70-3.56 (m, 5 H, H-5', H-6b, H-6', CH₂aCNH), 3.55-3.35 (m, 5 H, H-3, CH₂bCN, H-4, H-5, H-2'), 1.91 (s, 3 H, Ac); **¹³C NMR** (D₂O, 101 MHz) δ 174.71 (CO), 103.50 (C-1'), 83.60 (C-1), 83.05 (C-3), 79.70 (C-5'), 75.32 (C-4), 72.53 (C-5), 70.68 (C-2'), 68.57 (C-4'), 68.54 (CH₂CNH), 68.49 (C-3'), 61.07 (C-6), 60.77 (C-6'), 53.20 (C-2), 22.20 (Ac).

2-N-acetamido-4-O-[3-O-(N-acetyl- α -D-neuraminic acid)- β -D-galactopyranosyl]-1-thio-(S-2-imido-2-methoxyethyl)-2-deoxy- β -D-glucopyranoside **2.90**



To a solution of **2.73** (20 mg, 0.03 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 6 μ L, 0.03 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.90**) in ratio 1:0.9 (determined by NMR, ¹H, MeOD) , and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 500 MHz) δ 4.78 (d, *J* = 10.3 Hz, 1 H, H-1), 4.52 (d, *J* = 7.8 Hz, 1 H, H-1'), 4.08 (dd, *J* = 9.8, 2.5 Hz, 1 H, H-4), 4.03-3.90 (m, 2 H, H-6'', H-9a), 3.89-3.49 (m, 17 H, H-9b, CH₂CNH, H-4''', H-5''', H-7, H-8, H-2, H-3, H-5, H-6, H-2', H-3', H-4', H-5', H-6'), 3.31 (s, 3 H, OMe), 2.72 (dd, *J* = 12.3, 4.3 Hz, 1 H, H-3'''a), 2.00, 1.99 (2 x s, 2 x 3 H, 2 x Ac), 1.76 (t, *J* = 12.1 Hz, 1 H, H-3'''b); **¹³C NMR** (D₂O, 126 MHz) δ 175.07, 174.68, 173.91 (2 x CO, COOH), 118.57 (CNH), 102.62 (C-1'), 99.87 (C-2'''), 83.83 (C-1), 78.97, 77.95, 75.54, 75.24, 73.41, 72.94, 71.81, 69.39, 68.37, 68.17, 67.55, (C-4''', C-5''', C-8, CH₂CNH, C-3, C-4, C-5, C-2', C-3', C-4', C-5'), 62.64, 61.08, 60.14, 60.10 (C-6, C-6', C-6''', C-9), 53.91 (C-2), 51.76 (C-7), 48.92 (OMe), 39.71 (C-3'''), 22.14, 22.09 (2 x Ac).

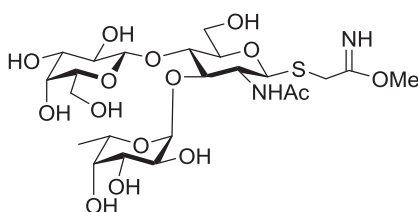
2-N-acetamido-3-O-(fucopyranosyl)-4-O-[3-O-(N-acetyl- α -D-neuraminic acid)- β -D-galactopyranosyl]-1-thio-(S-2-imido-2-methoxyethyl)-2-deoxy- β -D-glucopyranoside **2.91**



To a solution of **2.74** (20 mg, 0.02 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 5 μ L, 0.02 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.91**) in ratio 1:1 (determined by NMR, ¹H, MeOD) , and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 400 MHz) δ 5.10 (d, *J* = 4.0 Hz, 1 H, H-1''), 4.81 (d, *J* = 11.3 Hz, 1 H, H-1), 4.50 (d, *J* = 7.7 Hz, 1 H, H-1'), 4.10-3.42 (m, 25 H, H-2, H-3, H-4, H-5, H-6, H-2', H-3', H-4', H-5', H-6', H-2'', H-3'', H-4'', H-5'', H-4''', H-5''', H-6''', H-7, H-8, H-9, CH₂CNH), 3.32 (s, 3 H, OMe), 2.81-2.69 (m, 1 H, H-3'''a), 2.00 (s, 6 H, 2 x Ac), 1.76 (t, *J* = 12.3 Hz, 1 H, H-3'''b), 1.14 (d, *J* = 6.2 Hz, 3 H, H-6''); **¹³C NMR** (D₂O, 101 MHz) δ 175.10, 173.91 (2 x CO, COOH), 127.20 (CNH), 101.69 (C-1'), 99.73 (C-2'''), 98.53 (C-1''), 82.82, 82.72, 82.20, 80.79, 79.64, 79.48, 75.97, 75.48, 74.79, 72.91, 71.96, 71.27, 69.26, 67.84, 67.72, 66.53 (C-1, C-3, C-4, C-5, C-2', C-3', C-4', C-5', C-2'', C-3'', C-4'', 152

C-5'', C-4'', C-5'', CH₂CNH, C-8), 61.63, 61.54, 60.99, 60.69 (C-6, C-6', C-6'', C-9), 51.77 (C-2), 48.98 (C-7), 48.70 (OMe), 39.54 (C-3''), 22.11 (2 x Ac), 15.32 (C-6'').

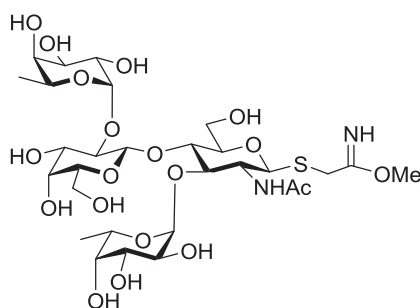
2-N-acetamido-3-O-(fucopyranosyl)-4-O-[β-D-galactopyranosyl]-1-thio-(S-2-imido-2-methoxyethyl)-2-deoxy-β-D-glucopyranoside **2.92**



To a solution of **2.24** (20 mg, 0.02 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 5 μL, 0.02 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.92**) in ratio 1:1 (determined by NMR, ¹H, MeOD) , and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 400 MHz) δ 5.10 (d, *J* = 3.7 Hz, 1 H, H-1''), 4.80 (d, *J* = 10.7 Hz, 1 H, H-1), 4.43 (d, *J* = 7.5 Hz, 1 H, H-1'), 4.10-3.80 (m, 8 H, H-2, H-5', H-5'', H-4, H-4', H-4'', H-6), 3.79-3.51 (m, 9 H, H-6', H-2'', CH₂CNH, H-5, H-3'', H-3', H-3), 3.51-3.41 (m, 1 H, H-2'), 1.99 (s, 3 H, Ac), 1.14 (d, *J* = 6.3 Hz, 3 H, H-6''); **¹³C NMR** (D₂O, 101 MHz) δ 174.42 (CO), 101.86 (C-1'), 98.78 (C-1''), 83.79 (C-1), 79.69, 76.12, 74.96, 73.40, 73.11, 72.52, 71.95, 71.07, 69.25, 68.39, 67.71, 66.79 (C-3, C-4, C-5, CH₂CNH, C-2', C-3', C-4', C-5',

C-2'', C-3'', C-4'', C-5''), 61.53 (C-6'), 59.83 (C-6), 54.49 (C-2), 22.22 (Ac), 15.34 (C-6'').

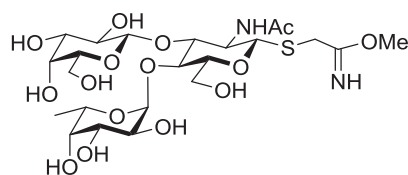
2-N-acetamido-3-O-(fucopyranosyl)-4-O-[2-O-(fucopyranosyl)- β -D-galactopyranosyl]-1-thio-(S-2-imido-2-methoxyethyl)-2-deoxy- β -D-glucopyranoside **2.93**



To a solution of **2.33** (20 mg, 0.02 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 4 μ L, 0.02 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.93**) in ratio 1:1 (determined by NMR, ¹H, MeOD), and used in the next step without further purification. **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (D₂O, 400 MHz) δ 5.24 (d, *J* = 1.6 Hz, 1 H, H-1'''), 5.10 (d, *J* = 4.0 Hz, 1 H, H-1''), 4.79 (d, *J* = 3.8 Hz, 1 H, H-1), 4.42 (d, *J* = 7.7 Hz, 1 H, H-1'), 4.21 (d, *J* = 6.3 Hz, 1 H, H-5''), 4.06-3.53 (m, 19 H, H-5''', CH₂CNH, H-2, H-3, H-4, H-5, H-6, H-3', H-4', H-5', H-6', H-2'', H-3'', H-4'', H-3''', H-4''', H-2'''), 3.52-3.41 (m, 1 H, H-2'), 3.31 (s, 3 H, OMe), 1.99 (s, 3 H, Ac), 1.22

(d, $J = 7.5$ Hz, 3 H, H-6''), 1.14 (d, $J = 6.1$ Hz, 3 H, H-6'''); **^{13}C NMR** (D_2O , 101 MHz) δ 174.44 (CO), 137.94 (CNH), 101.87 (C-1'), 100.31 (C-1'''), 98.78 (C-1''), 83.80 (C-1), 79.97, 79.69, 76.49, 76.04, 74.96, 73.54, 73.12, 72.53, 71.95, 71.08, 70.52, 69.76, 69.25, 68.39, 67.71 (C-2, C-3, C-4, C-5', C-5'', C-2', C-3', C-4', CH_2CNH , C-2'', C-3'', C-4'', C-2''', C-3''', C-4'''), 66.85 (C-5), 61.53 (C-6), 61.36 (C-6'), 59.81 (C-5'''), 43.65 (OMe), 22.22 (Ac), 15.49 (C-6''), 15.34 (C-6''').

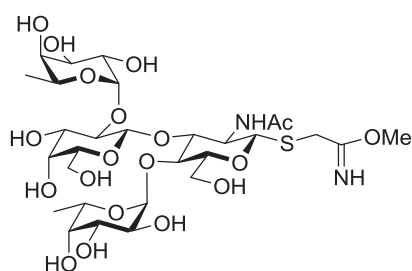
2-N-acetamido-4-O-(fucopyranosyl)-3-O-[β -D-galactopyranosyl]-1-thio-(S-2-imido-2-methoxyethyl)-2-deoxy- β -D-glucopyranoside **2.94**



To a solution of **2.29** (20 mg, 0.02 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 5 μL , 0.02 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.94**) in ratio 1:1 (determined by NMR, ^1H , MeOD), and used in the next step without further purification. **TLC** $R_f = 0.0$ (DCM:MeOH, 10:1); **^1H NMR** (D_2O , 400 MHz) δ 5.12 (d, $J = 2.6$ Hz, 1 H, H-1''), 4.98 (d, $J = 3.7$ Hz, 1 H, H-1), 4.63 (d, $J = 7.5$ Hz, 1 H, H-1'), 4.27 (d, $J = 6.7$ Hz, 1 H, H-5''), 4.14 (dd, $J = 18.6, 9.2$ Hz, 1 H, H-4'), 4.01-3.66 (m, 14 H, H-5', H-2, H-4'', H-2'', H-3', H-3, H-4, H-6, H-4''', H-6', H-2''', H-3'''), 3.57 (dd, $J =$

9.9, 6.9 Hz, 2 H, H-5, H-2'), 2.03 (s, 3 H, Ac), 1.22 (d, $J = 6.5$ Hz, 5 H, H-6''); ^{13}C NMR (D_2O , 101 MHz) δ 173.55 (CO), 100.69 (C-1'), 99.59 (C-1''), 97.89 (C-1), 79.97, 76.47, 75.89, 74.89, 73.72, 72.03, 69.48, 68.81, 68.32, 67.85, 67.10, 66.27 (C-3, C-4, C-5, CH_2CNH , C-2', C-3', C-4', C-5', C-2'', C-3'', C-4'', C-5''), 61.63 (C-6'), 59.63 (C-6), 54.23 (C-2), 22.17 (Ac), 15.29 (C-6'').

2-N-acetamido-4-O-(fucopyranosyl)-3-O-[2-O-(fucopyranosyl)- β -D-galactopyranosyl]-1-thio-(S-2-imido-2-methoxyethyl)-2-deoxy- β -D-glucopyranoside **2.95**



To a solution of **2.39** (20 mg, 0.02 mmol, 1 eq.) in MeOH (dry, 1 mL) was added NaOMe (25% in MeOH, 4 μL , 0.02 mmol, 1 eq.). The reaction was stirred at r.t. for 16 h before being neutralised with DowexH⁺. The reaction mixture was filtered and evaporated without heating. The product was obtained as a mixture of deprotected (R-SCM) and activated sugars (R-IME, **2.95**) in ratio 1:1 (determined by NMR, ^1H , MeOD), and used in the next step without further purification. TLC $R_f = 0.0$ (DCM:MeOH, 10:1); ^1H NMR (D_2O , 500 MHz) δ 5.12 (d, $J = 2.6$ Hz, 1 H, H-1'''), 4.98 (d, $J = 3.4$ Hz, 1 H, H-1''), 4.80 (d, $J = 6.9$ Hz, 1 H, H-1), 4.63 (d, $J = 7.4$ Hz, 1 H, H-1'), 4.27 (d, $J = 6.6$ Hz, 1 H, H-5''), 4.15 (t, $J = 9.5$ Hz, 1 H, H-4'), 4.02-3.63 (m, 17 H, H-5''', CH_2CNH , H-3, H-2, H-4, H-6, H-3', H-5', H-6', H-2'', H-

156

3'', H-4'', H-2''', H-3''', H-4'''), 3.57 (dd, $J = 10.3, 6.4$ Hz, 2 H, H-5, H-2'), 2.03 (s, 3 H, Ac), 1.22 (d, $J = 6.4$ Hz, 6 H, H-6'', H-6'''); ^{13}C NMR (D_2O , 101 MHz) δ 173.55 (CO), 100.39 (C-1'), 99.57 (C-1'''), 97.89 (C-1''), 79.89 (C-1), 76.15, 75.80, 75.13, 74.88, 73.72, 71.99, 71.89, 70.78, 69.51, 68.80, 67.84, 67.64, 67.00, 66.64, 66.34 (C-2, C-3, C-4, C-5', C-5'', C-2', C-3', C-4', CH_2CNH , C-2'', C-3'', C-4'', C-2''', C-3''', C-4'''), 65.99 (C-5), 61.62 (C-6), 60.70 (C-6'), 59.42 (C-5'''), 45.83 (OMe), 22.16 (Ac), 15.29 (C-6''), 14.65 (C-6''').

BCA assay

BCA mix was prepared by mixing 3 mL of solution A and 60 μL of solution B. A calibration curve (Fig. S1) was obtained from:

10 μL of standard solution of BSA (0-2 mg/mL) + 200 μL of BCA mix, incubated at 37 $^\circ\text{C}$ for 30 min (the absorbance (Abs) read at 562 nm).

Conc	Abs
0.000	0.0800
0.050	0.1133
0.075	0.1241
0.100	0.1395
0.125	0.1569
0.250	0.2377
0.500	0.3626
0.750	0.4804
1.000	0.5985
2.000	0.9958

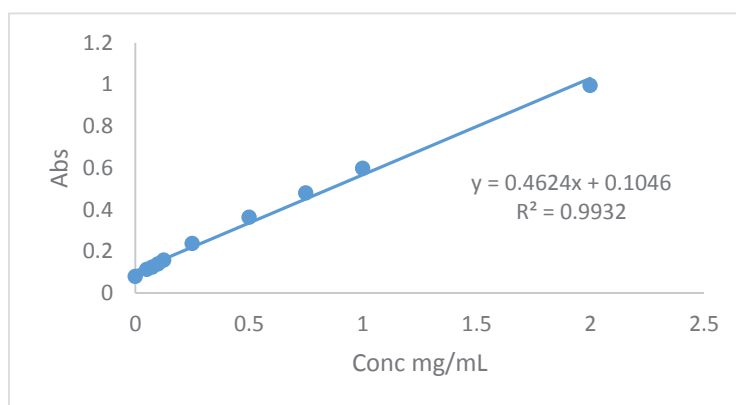


Figure S1. BCA calibration curve from the Abs measured for BSA at progressive dilutions.

Protein expression

A starter culture of PmST was used to inoculate LB media (1000 mL), supplemented with ampicillin (100 µg/mL), the cells incubated at 37 °C, and the optical density (OD₆₀₀) monitored until a value of 0.6-0.8 was obtained (~3-4 h). Expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG, 1 mM) and was continued for 16 h at 30 °C with shaking (220 rpm). The cells were harvested by centrifugation (8 min, 8,000 rpm) to give pellets, which were collected in a 50 mL falcon tube and re-suspended in binding buffer (Tris-HCl 20 mM, imidazole 15 mM, NaCl 0.5 M, 1% glycerol, pH 7.8, 10 mL). The cell suspension was treated with DNase (10 µL), and proteinase inhibitor tablet (1/4 tablet), and rocked at 4 °C for 2 h. The cells were then lysed by sonication (Soniprep at 10 micron amplitude, 30 sec on with 30 sec rest, 5 cycles, on ice). The cell debris was removed from the lysis mixture by centrifugation (30 min, JA 25.5 rotor, 20,000 rpm). The protein was purified by incubation of the cleared cell lysate with HisBind Ni-NTA resin (3 mL per litre of culture, Novagen) for 2 h at 4 °C. The resin was packed into a glass econo-column and unbound proteins washed out with binding buffer (A) (Tris-HCl 20 mM, imidazole 15 mM, NaCl 0.5 M, 1% glycerol, pH 7.8) (15-20 CV, ~50 mL). The bound protein was eluted from the resin by elution buffer (B) (Tris-HCl 20 mM, imidazole 500 mM, NaCl 0.5 M, 5% glycerol, pH 7.8) using a step gradient (20% B (5 mL), 40% B (5 mL), 60% B (5 mL), 80% B (5 mL), 100% B (4 x 5 mL)). All eluted proteins were collected in 5 mL fractions, which were analysed by SDS-PAGE (Fig.S2).

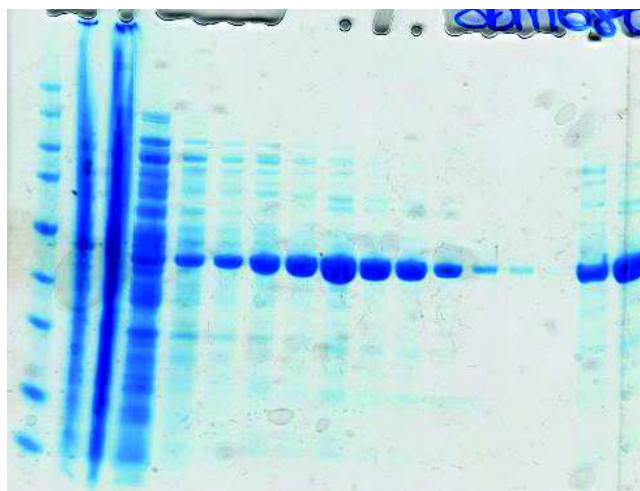


Figure S2. SDS-PAGE gel of PmST purification. From the left: ladder (first) and elution fractions. Single spots are considered pure protein.

Fractions containing pure PmST protein were combined, subjected to ammonium sulphate precipitation (85% saturation) for 16 h at 4 °C. The precipitated proteins were pelleted by centrifugation (40 minutes, JA 25.5 rotor, 20000 rpm, 4 °C), desalted by overnight dialysis against sodium phosphate buffer (PB) (50 mM, pH 7.5) then analysed by LC-MS.

Protein concentration for PmST was determined by BCA assay (4.19 mg/mL).

The protein was then analysed by LC-MS (Calc: 46411, Found: 46415, Fig. S3)

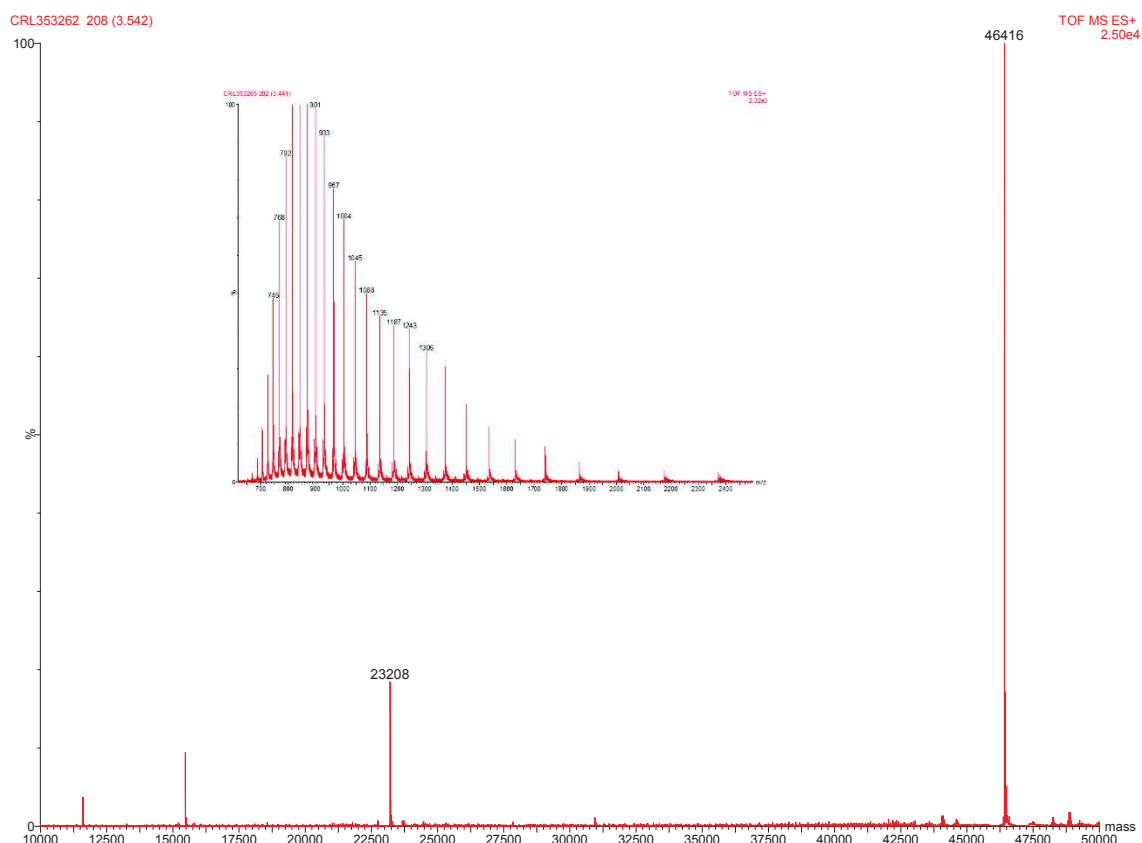


Figure S3. ESI-MS of pure PmST. Superimposed is the ion series before deconvolution.

Starter culture of NmCSS was used to inoculate LB media (1000 mL), supplemented with ampicillin (100 µg/mL), the cells incubated at 37 °C, and the optical density (OD₆₀₀) monitored until a value of 0.6-0.8 was obtained (~3-4 h). Expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG, 1 mM) and was continued for 16 h at 30 °C with shaking (220 rpm). The cells were harvested by centrifugation (8 min, 8,000 rpm) to give pellets, which were collected in a 50 mL falcon tube and re-suspended in binding buffer (Tris-HCl 20 mM, imidazole 15 mM, NaCl 0.5 M, 1% glycerol, pH 7.8, 10 mL). The cell suspension was treated with DNase (10 µL), and proteinase inhibitor tablet (1/4 tablet), and rocked at 4 °C for 2 h. The cells were then lysed by sonication

(Soniprep at 10 micron amplitude, 30 sec on with 30 sec rest, 5 cycles, on ice). The cell debris was removed from the lysis mixture by centrifugation (30 min, JA 25.5 rotor, 20,000 rpm). The protein was purified by incubation of the cleared cell lysate with HisBind Ni-NTA resin (3 mL per litre of culture, Novagen) for 2 h at 4 °C. The resin was packed into a glass econo-column and unbound proteins washed out with binding buffer (A) (Tris-HCl 20 mM, imidazole 15 mM, NaCl 0.5 M, 1% glycerol, pH 7.8) (15-20 CV, ~50 mL). The bound protein was eluted from the resin by elution buffer (B) (Tris-HCl 20 mM, imidazole 500 mM, NaCl 0.5 M, 5% glycerol, pH 7.8) using a step gradient (20% B (5 mL), 40% B (5 mL), 60% B (5 mL), 80% B (5 mL), 100% B (4 x 5 mL)). All eluted proteins were collected in 5 mL fractions, which were analysed by SDS-PAGE (Fig.S4).

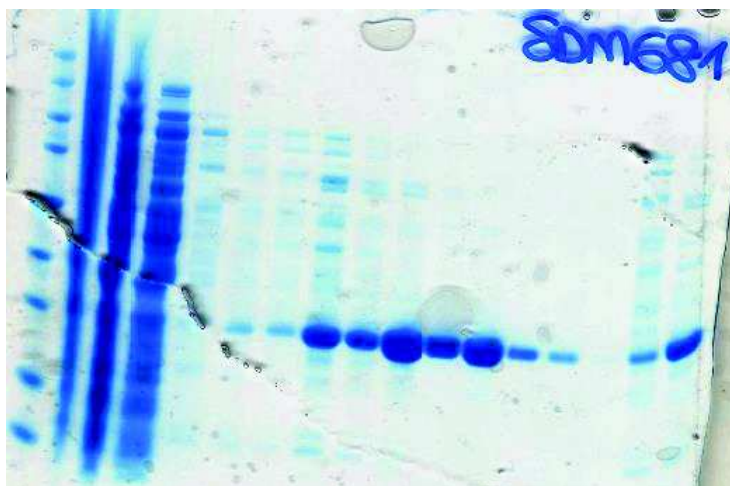


Figure S4. SDS-PAGE gel of NmCSS purification. From the left: ladder (first) and elution fractions. Single spots are considered pure protein.

Fractions containing pure NmCSS protein were combined, subjected to ammonium sulphate precipitation (85% saturation) for 16 h at 4 °C. The precipitated proteins were pelleted by centrifugation (40 minutes, JA 25.5 rotor,

20000 rpm, 4 °C), desalted by overnight dialysis against sodium phosphate buffer (PB) (50 mM, pH 7.5) then analysed by LC-MS.

Protein concentration for NmCSS was determined by BCA assay (3.97 mg/mL).

The protein was then analysed by LC-MS (Calc: 26,089, Found: 26,090, Fig. S5)

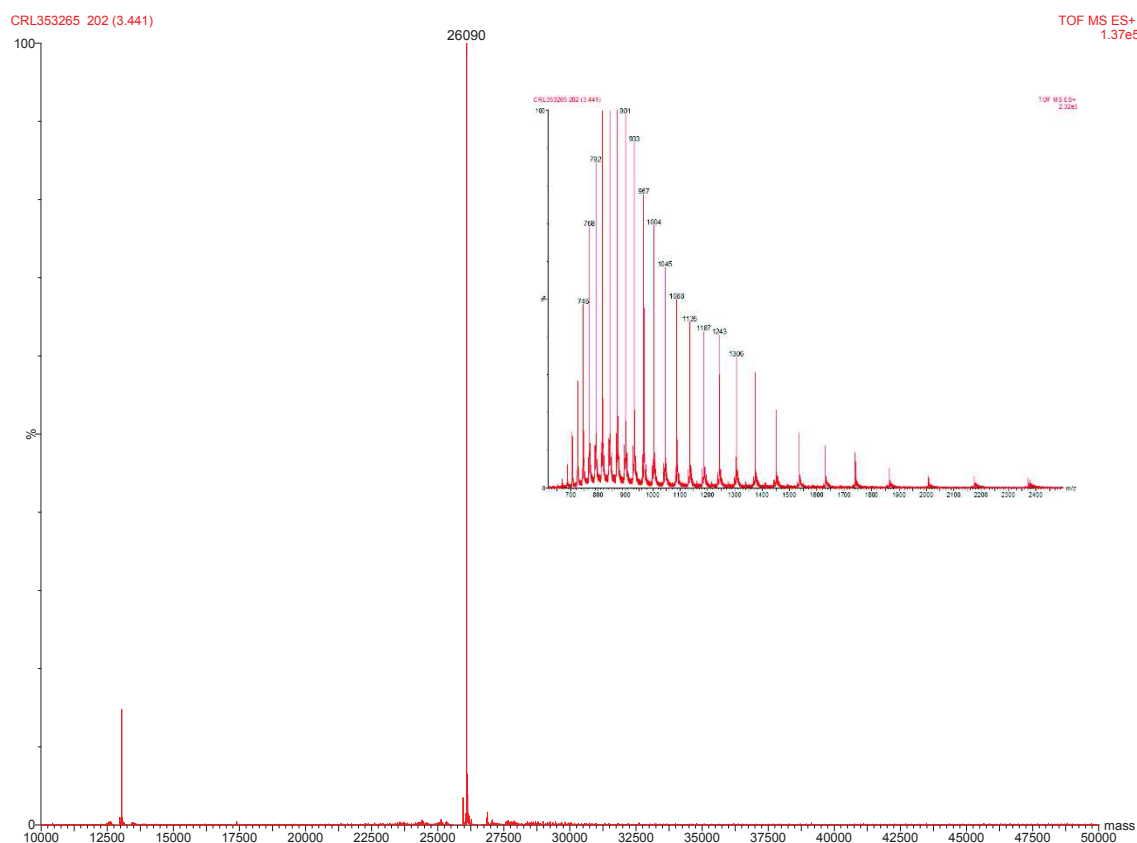


Figure S5. ESI-MS of pure NmCSS. Superimposed is the ion series before deconvolution.

References

- [1] B. R. Baker, J. P. Joseph, and J. H. Williams, *J. Org. Chem.* **1954**, 19, 1793-1801.
- [2] Y. Inouye, K. Onodera, S. Kitaoka, and H. Ochiai, *J. Am. Chem. Soc.* **1957**, 79, 4218-22.
- [3] G. J. L. Bernardes, E. J. Grayson, S. Thompson, J. M. Chalker, J. C. Errey, F. E. Oualid, T. D. W. Claridge, and B. G. Davis, *Angew. Chemie Int. Ed.* **2008**, 47, 2244-7.
- [4] Y. C. Lee, C. P. Stowell, and M. J. Krantz, *Biochemistry* **1976**, 15, 3956-63.
- [5] S. I. van Kasteren, S. J. Campbell, S. Serres, D. C. Anthony, N. R. Sibson, and B. G. Davis, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, 106, 18-23.
- [6] W. Haskins, M. Raymond, M. Hann, and C. Hudson, *J. Am. Chem. Soc.* **1942**, 64, 1852-6.
- [7] E. Fischer and E. Pischer, *Helv. Chim. Acta* **1928**, 11, 910-5.
- [8] R. R. Schmidt, *Liebigs Ann. Chem.* **1983**, 1249-56.
- [9] R. Daly, G. Vaz, A. M. Davies, M. O. Senge, and E. M. Scanlan, *Chem. Eur. J.* **2012**, 18, 14671-9.
- [10] H. Paulsen, F. Garrido, W-P. Trautwein, *Chem. Ber.* **1968**, 101, 186-90.
- [11] H. J. Koeners, J. Verhoeven, and J. H. van Boom, *Tetrahedron Lett.* **1980**, 21, 381-2.
- [12] H. J. Koeners, J. Verhoeven, J. H. van Boom, *Journal of the Royal Netherlands Chemical Society* **1981**, 100, 65-72.

- [13] H. J. Koeners, C. Schattenkerk, J. Verhoeven, J. H. van Boom, *Tetrahedron Lett.* **1980**, 21, 2373-6.
- [14] L. M. Lerner, *Carbohydr. Res.* **1971**, 19, 255-8.
- [15] E. J. Behrman, *Carbohydr. Res.* **1978**, 64, 297-301.
- [16] Z. Zhiyuan and G. Magnusson, *Carbohydr. Res.* **1994**, 262, 79-101.
- [17] J. A. F. Joosten, B. Evers, R. P. van Summeren, J. P. Kamerling, and J. F. G. Vliegthart, *European J. Org. Chem.* **2003**, 2003, 3569-86.
- [18] P. A. Levene, I. Bencowitz, *J. Biol. Chem.* **1927**, 74, 153-74.
- [19] G. R. Morais, A. J. Humphrey, and R. A. Falconer, *Carbohydr. Res.* **2009**, 344, 1039-45.
- [20] O. M. T. Pearce, K. D. Fisher, J. Humphries, L. W. Seymour, A. Smith, and B. G. Davis, *Angew. Chem. Int. Ed. Engl.* **2005**, 44, 1057-61.
- [21] Z. Tarasiejska and R. W. Jeanloz, *J. Am. Chem. Soc.* **1958**, 80, 6325-7.
- [22] B. S. Patil and V. V. S. Babu, *Indian J. Chem.* **2004**, 43, 1288-91.
- [23] M. Bergmann, *Justus Liebigs Ann. Chem.* **1923**, 434, 79-110.
- [24] Z. Zhang and G. Magnusson, *Carbohydr. Res.* **1996**, 925, 41-55.
- [25] W. Korytnyk, J. A. Mills, *Journal of Biological Chemistry* **1924**, 59, 126.
- [26] *J Chem Soc* **1959**, 636-49.
- [27] J. K. Dale, *J. Am. Chem. Soc.* **1924**, 46, 1046-51.
- [28] *Journal of Biological Chemistry* **1931**, 90, p. 89-95.
- [29] S. Tejima, T. Maki, and M. Akagi, *Chem. Pharm. Bull.* **1964**, 12, 528-32.
- [30] I. Pirmettis, Y. Arano, T. Tsotakos, K. Okada, A. Yamaguchi, T. Uehara, M. Morais, J. D. G. Correia, I. Santos, M. Martins, S. Pereira, C. Triantis, P. Kyprianidou, M. Pelecanou, and M. Papadopoulos, *Mol. Pharm.* **2012**, 9, 1681-92.

- [31] *Justus Liebigs Annalen der Chemie* **1912**, 394, 225.
- [32] J. L. Chen, R. E. Moore, and G. M. L. Patterson, *J. Org. Chem.* **1991**, 56, 4360-64.
- [33] H. von Kurt and H. Scharman, *Liebigs Ann. Chem. Bd.* **1963**, 667, 183-93.
- [34] D. H. Brauns, *J. Am. Chem. Soc.* **1923**, 45, 833-5.
- [35] R. J. Ferrier, P. C. Tyler, *J. Chem. Soc. Perkin Trans. 1* **1980**, 2767-73.
- [36] O. Chemie and D. U. Rostock, *Chem. Ber.* **1960**, 2705-12.
- [37] M. Ďuranova, J. Hirsch, K. Kolenova and P. Biely, *Biosci. Biotechnol. Biochem.* **2009**, 73, 2483-7.
- [38] F. Goebel, *J. Biol. Chem.* **1935**, 111, 347-53.
- [39] W. Bowering, *J. Am. Chem. Soc.* **1960**, 82, 2827-30.
- [40] P. Press, *Tetrahedron Lett.* **1967**, 44, 4363-9.
- [41] B. Bressel and H. Reissig, *Org. Lett.* **2009**, 11, 527-30.
- [42] S. B. Cohen and R. L. Halcomb, *J. Org. Chem.* **2000**, 65, 6145-52.

A triply divergent reagent for glycoprotein synthesis

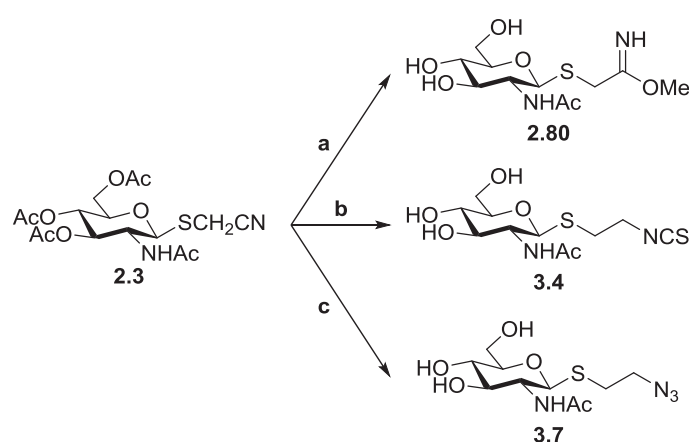
3.1 Introduction

3.1.1 Synthesis of glycoproteins by a triply divergent reagent

The efforts spent towards the synthesis of glycoproteins containing well-defined glycosylation patterns have been escalating during the last decades.^[1-3] The reasons that motivate such interest come from the basic research of their biological roles^[4] (e.g. antibodies, receptors, enzymes) as well as from the marketing of biotherapeutics^[5] (e.g. vaccines, hormones). The aim is to design protein structures bearing a key triad of features: unambiguous glycan structures, tight control over carbohydrate copy number, and high site-specificity for protein modifications. These requirements are essential to understand how small differences in glycosylation patterns impart dramatically different functions to similar structures.^[6,7] The decoration of proteins with glycans can be achieved both enzymatically and chemically by reacting pre-activated carbohydrates^[1-3] with amino acids possessing reactive functional groups.

Although our achievements in expressing and modifying natural and unnatural proteins are partially responsible for the evolution of our knowledge of their biological roles, we still possess limited tools to access very complex structures.

Using the experience we gained in manipulating carbohydrates we proposed to approach this formidable task by looking at the synthesis of glycoproteins from the carbohydrates' perspective. From the diversification of a single sugar reagent we tackled the glycosylation of different protein scaffolds. Here, from the precursor **2.3**, we generated three glycosylating reagents in a triply divergent manner^[8] (Scheme 3.1).



Scheme 3.1. Reagents and conditions: a) NaOMe, MeOH, r.t., 50%; b) (i) H₂, PtO₂, EtOH:CHCl₃ (5:1), r.t.; (ii) NaOMe, MeOH, r.t.; (iii) thiophosgene, Na₂CO₃:NaHCO₃ (1:1, 0.3 M), r.t., quant.; c) (i) H₂, PtO₂, EtOH:CHCl₃ (5:1), r.t.; (ii) **3.5**, MeOH, CuSO₄, r.t.; (iii) NaOMe, MeOH, r.t., quant..

While **2.80** can be synthesised from the common precursor **2.3** in one single step, **3.4** and **3.7** syntheses proceed *via* an amino intermediate. The advantage of this strategy is to open up multiple options for the glycosylation of a protein without having to manipulate the protein on a macromolecular level but instead acting on the more pliable saccharide unit.

3.1.2 Site-specific protein modification

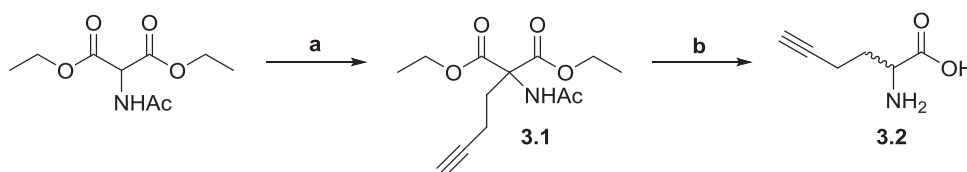
Selectively modifying a specific site (or amino acid) in a protein may affect its folding, activity, binding and more. Modifications of key amino acids in active sites may result in loss of activity, while derivatisation of remote residues may result in less drastic outcomes. After the identification of those residues that would be suitable candidates for the decoration of a protein there are several methodologies available in our toolbox^[9] to introduce such modifications: from repositioning of low abundant amino acids to the ‘tag-and-modify’ approach.^[10] Some natural amino acids bear reactive functionalities that may be utilised in bio-conjugation reactions. Probably the most representative examples are cysteine (Cys) and lysine (Lys) residues.^[11,12] This approach holds great potential, but has some substantial limitations. Amongst them is the limited chemistry available on thiol and amine groups, which have to be compatible with mild conditions and an aqueous environment necessary for protein stability.

More attention has been drawn by the ‘tag-and-modify’ approach, which consists of installing an unnatural amino acid (tag) in a site-specific manner in the protein which will subsequently function as a chemoselective access point for the planned modification. This approach has attracted more interest as it makes available a larger number of chemical residues - and therefore reactions - to chemical biologists, such as thiol-ene couplings,^[13-14] Suzuki-Miyaura couplings,^[15-17] Staudinger ligations,^[18] azide-alkyne cycloadditions^[19,20] and cross metathesis.^[21]

3.2 Protein scaffolds

3.2.1 Homopropargylglycine (Hpg)

To modify protein candidates in both a site-specific and non-site-specific manner we decided to use the ‘tag-and-modify’ approach. An unnatural amino acid (Hpg) was introduced into their structures, replacing a Met in a defined position in their sequence, allowing for subsequent site-specific modification. Hpg possesses a terminal alkyne on its side chain (Scheme 3.2), which we predict will react with our azido-sugar derivative in a copper catalysed azide-alkyne cycloaddition (CuAAC).



Scheme 3.2. Reagents and conditions: a) (i) NaH, DMF, 0 °C; (ii) 4-bromo butyne, 90 °C, 12%; b) (i) NaOH, MeOH:H₂O (1:1), 100 °C; (ii) HCl 2 M, 100 °C, 44%.

The proteins we selected as scaffolds were Np276 and Q β (described below). The reason for our choice was to select structurally and biologically distinct proteins, having a different degree of complexity, where Hpg could be selectively introduced. As we will shortly describe, Np276 is a monomeric helical protein while Q β is a more complex 180-mer protein which self-assembles into a virus-like particle.

3.2.2 Np276-M61Hpg

Np276 (Fig. 3.1) is a highly symmetric right-handed quadrilateral β -helix protein of 200 amino acids from the genome of *Nostoc punctiforme*,^[22] a species of filamentous cyanobacteria.

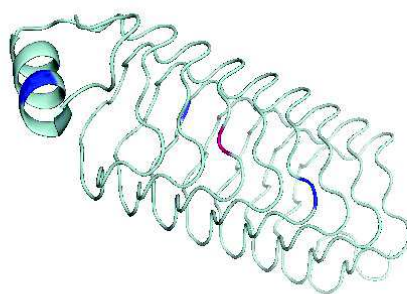


Figure 3.1. Np276-M61Hpg structure. In red Hpg and in blue Lys residues.

We reassigned the Met sense codon to introduce the unnatural amino acid Hpg in position 61:

```
PSSHHHHHHSSGLVPRGSHIDVGKLRQLYAAGERDFS  
IVDLRGAVLENINLSGAILHGAMLDEANLQQANLSRAD  
LSGATLNGADLRGANLSKADLSDAILDNAILEGAILDEA  
VLNQANLKAANLEQAILSHANIREADLSEANLEAADLS  
GADLAIADLHQANLHQAAALERANLTGANLEDANLEGTI  
LEGGNNNLAT
```

This protein was expressed in the *E. coli* B834(DE3) methionine auxotrophic strain and purified by His-tag affinity purification. The incorporation of Hpg was

found to be >95% in all cases. The final Np276-M61Hpg^[23] contained 1 Hpg (Fig. 3.1 in red) and 3 Lys (Fig. 3.1 in blue) residues, ready to react with our glycosylating reagents.

3.2.3 Q β -M16Hpg

Q β (Fig. 3.2) is a coat protein from *Allolevivirus Q-beta*^[24] made of 180-copies of the 132 amino acid long monomer, which self assembles to form an icosahedral nanoparticle. Disulphide bonds between the monomers stabilise the construct, although they have been proven not to be essential for its assembly.^[25]

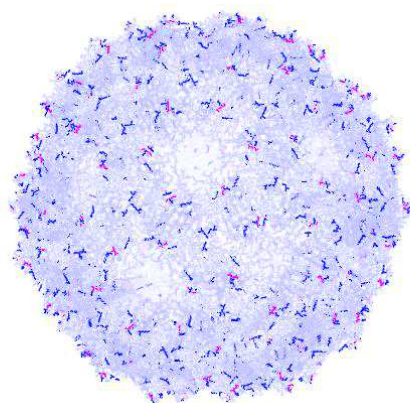


Figure 3.2. Q β -M16Hpg structure. In red Hpg and in blue Lys residues.

As for Np276, the Met sense codon was reassigned to introduce the unnatural amino acid Hpg in position 16:

```
AKLETVTTLGNIGKDGMQTLVLNPRGVNPTNGVASLSQA  
GAVPALEKRVTVSVSQPSRNRKNYKVQVKIQNPTACT  
ANGSCDPSVTRQAYADVTFSTQYSTDEERAFVRTTEL  
AALLASPLLIDAIDQLNPAY
```

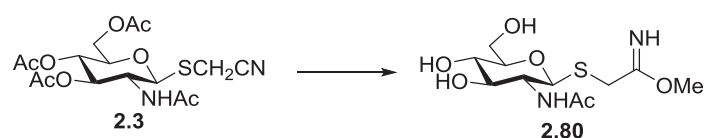
Here again we used the *E. coli* B834(DE3) methionine auxotroph, to achieve >95% incorporation of Hpg. The protein was isolated as a mixture of monomers, multimers and intact nanoparticles, which were purified by sucrose gradient. The crude mixture was deposited on a sucrose gradient of increasing concentration and subjected to ultracentrifugation. The density difference between the layers separated components of different size/weight, allowing for the isolation of pure and fully formed nanoparticles.

The final Q β -M16Hpg^[26] therefore contained 1 x 180 Hpg (Fig. 3.2 in red) and 6 x 180 Lys (Fig. 3.2 in blue), ready to be glycosylated at a later stage.

3.3 A triply divergent reagent

3.3.1 From CN to IME

Synthesis of the IME reagent^[27] (**2.80**, Scheme 3.3) from **2.3** has already been described in the previous chapter (See chapter 2.7).



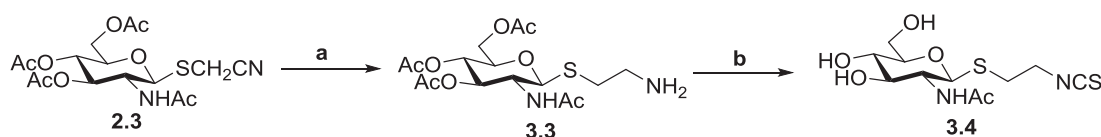
Scheme 3.3. Reagents and conditions: NaOMe, MeOH, r.t., 50%.

2.3 was simultaneously deprotected and activated into IME by treating the starting material with NaOMe in MeOH^[28] overnight. Although the acetyl groups were removed in a matter of a few minutes, the reaction was carried on for 16 h as the formation of the IME is an equilibrium (see chapter 2.4) that takes longer

to equilibrate, resulting in around a maximum of 50% conversion. The imidate formed is expected to react promptly with primary amines on side-chains of Lys residues on both our protein scaffolds, in buffer and under mild conditions, generating an amidine^[29] linkage.

3.3.2 From CN to NCS

The synthesis of the NCS reagent^[30] (**3.4**, Scheme 3.5) from **2.3** goes through the amino intermediate **3.3**.



Scheme 3.5. Reagents and conditions: a) H₂, PtO₂, EtOH:CHCl₃ (5:1), r.t., 46%; b) (i) NaOMe, MeOH, r.t.; (ii) thiophosgene, Na₂CO₃:NaHCO₃ (1:1, 0.3 M), r.t., quant..

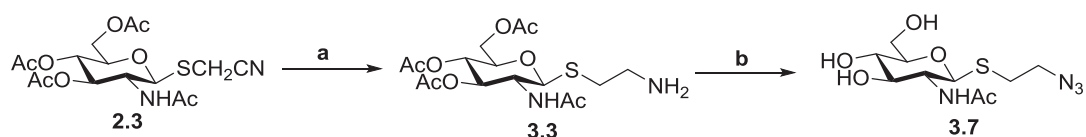
3.3 was obtained by reducing the cyano group into the primary amine by hydrogenation^[31] of the substrate. PtO₂ was used as a catalyst instead of the cheaper and more common Pd variants, as Pt is less sensitive to poisoning by the thioether moiety of the sugar.^[32] EtOH proved optimal for this reaction, using CHCl₃ as co-solvent to solubilise the starting material. Extremely dry conditions were unnecessary but air (oxygen) exclusion was imperative due to the nature of the reduction reaction. These mild conditions allowed a maximum conversion of almost 50%, with multiple additions of the catalyst and extremely lengthy reaction time (few days). Other conditions could have been applied to our system (e.g. use of different catalysts, acids etc.), however we anticipate that our conditions would allow the modification of more complex and acid-sensitive sugars (e.g.

carbohydrates such as Lewis saccharides, bearing the acid-sensitive fucose linkage).

Intermediate **3.3** was promptly deprotected in Zemplén conditions^[28] and then fully converted into the isothiocyanate^[30] derivative by thiophosgene.

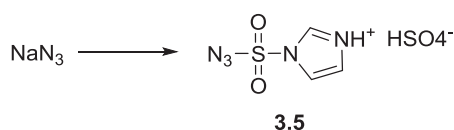
3.3.3 From CN to N₃

The synthesis of the N₃ reagent (**3.7**, Scheme 3.6) from **2.3** goes also through the amino intermediate **3.3**.



Scheme 3.6. Reagents and conditions: a) H₂, PtO₂, EtOH:CHCl₃ (5:1), r.t., 46%; b) (i) **3.5**, MeOH, CuSO₄, r.t.; (ii) NaOMe, MeOH, r.t., quant..

3.3 was converted into the azido compound **3.7** by the diazo-transfer reagent **3.5**. The synthesis of this reagent (Scheme 3.7) was based on studies conducted by the Klapötke group on the stability of imidazole-1-sulphonyl azide salts.^[33] We selected this particular compound to preserve the electron-withdrawing capacity of the sulphonyl moiety which governs the susceptibility of the diazo-transfer.



Scheme 3.7. Reagents and conditions: (i) SOCl₂, ACN, r.t.; (ii) imidazole, H₂SO₄, r.t., 53%.

Compound **3.5** proved to be stable for months if stored at -20 °C.

A quick deprotection of azido-sugar with NaOMe^[28] afforded the final compound **3.7** in quantitative yield, ready to react with Hpg terminal alkyne.

3.4 Protein glycosylation

3.4.1 Protein glycosylation via IME

The GlcNAc-imidate **2.80** was coupled on both Np276-M61Hpg and Q β -M16Hpg in basic aqueous conditions (pH 9), to preserve the reactivity of the deprotonated Lys.^[27] In our case several buffers, reaction times and equivalents of **2.80** were tested to achieve the full glycosylation of both proteins. While the glycosylation occurred non-specifically on all Lys residues for both proteins, the conversions obtained were quite different. For Q β , complete glycosylation could not be achieved, as some unreacted protein was always detectable even using up to 1000 equivalents of the sugar reagent and letting the reaction run for 16 h. Maximum conversion was reached in 1 h, using phosphate buffer (PB) (50 mM). Np276 represented a far easier case, as with the same equivalents (but using 0.6 M TAPS, for 2 h) we obtained di- and tri-glycosylated constructs. The resistance offered by Q β to complete conversion could be attributed to the fact that by reacting the intact virus-like particle in non-denaturing conditions, some Lys residues might be buried inside the folded structure and hence be less accessible to react.

3.4.2 Protein glycosylation via NCS

GlcNAc-isothiocyanate^[30] **3.4** was reacted in NaHCO₃ (0.3 M) overnight, for both proteins. This sugar reagent proved to be slower to react with Lys residues than the cognate imidate, but still afforded non-specific glycosylation patterns. Here again, while Np276 could be fully converted in to the more modest mono-GlcNAcylated derivative, traces of unreacted protein were still detected for the Q β multimer.

3.4.3 Protein glycosylation via N₃

GlcNAc-azide **3.7** reacted with Np276-M61Hpg and Q β -M16Hpg in a site-specific manner, *via* CuAAC.^[34] The copper catalyst was added to the reaction mixture as Cu(I)-TTMA complex,^[19] pre-formed in ACN from Cu(I)Br and TTMA^[35] (Fig.3.3).

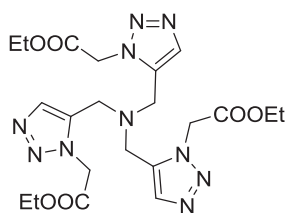


Figure 3.3. Tris(1-((O-ethyl)carboxymethyl)-1-H-1,2,3-triazol-4-ylmethyl) amine (TTMA), ligand forming the Cu(I)·TTMA complex.

Np276 reached full conversion yielding the mono-substituted product in just 3 h with only 3 additions of the catalyst complex, while Q β , again less prone to react, took more than 16 h and up to 5 additions of activated catalyst, but successfully reached complete conversion as well.

3.5 Binding experiments

By glycosylating Np276-M61Hpg and Q β -M16Hpg with our three sugar reagents, we introduced three different linkage types: amidine,^[29] thiourea^[30] and triazole.^[19,20] In order to evaluate how the influence of the linkage (pK_a , length, etc.) affects the functionality of the proteins, we studied their binding to the model lectin Wheat Germ Agglutinin (WGA).^[36] This carbohydrate-binding protein has an innate affinity for GlcNAc and therefore would recognise and bind GlcNAc-modified structures. The binding between GlcNAcylated proteins and WGA was studied by Surface Plasmon Resonance (SPR). This technique is based on the adsorption of one of the binding partners, in this case WGA, on a sensor chip. The chip is irradiated with a light beam, which interacts with the surface of the chip and is re-emitted at a specific angle. This angle changes when the other binding partner interacts with the immobilised WGA. Our samples of modified and un-modified proteins were flushed on the surface of the chip and the signal was recorded to detect when binding occurred (Fig. 3.4).

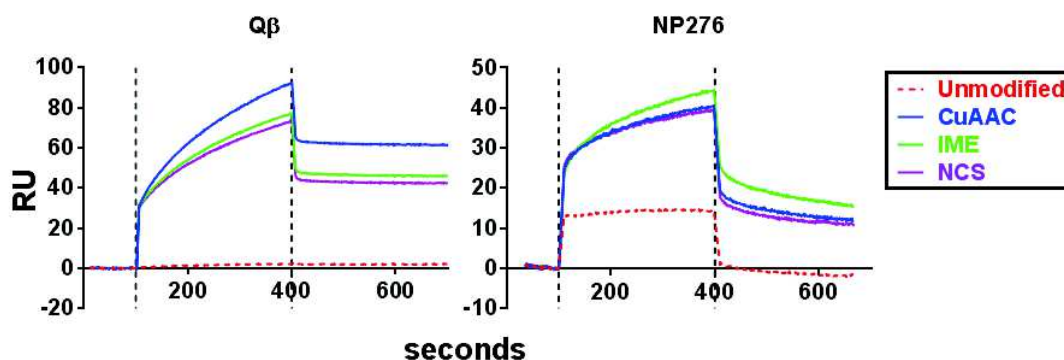


Figure 3.4. SPR binding of synthetic glycoproteins to the lectin WGA. Higher binding is shown for proteins bearing multiple sugar units (Q β modified by CuAAC and Np276 modified by IME couplings). No binding (or aspecific binding) is observed for non-glycosylated proteins.

Np276 glycosylated derivatives bound to WGA reflecting their degree of modification: when more sugars were presented by the protein the intensity of the binding was seen to increase. From the results obtained, no significance was attributable to the glycosylation site on the protein. Non glycosylated Np276 showed some non-specific binding, although the signal detected was of much lower intensity with respect to the signal given by the glycosylated analogues.

Glycosylated Q β particles showed stronger binding than Np276, possibly due to the higher glycan loading of each protein interacting with WGA. Here again, the intensity of their binding reflected their valency. No non-specific binding due to the unmodified Q β was detected.

By SPR we were able to detect the very weak interactions between glycosylated proteins and WGA, which could not observe by Western Blot.

3.6 Conclusions

Within this project we developed three different glycosylating reagents: a nucleophile, an electrophile and a 1,3-dipole from a single nitrile precursor. With these reagents we successfully achieved the glycosylation of two protein scaffolds of different complexity in both a site-specific and non-site-specific manner. The glycosylation occurred on natural (Lys) as well as on unnatural (Hpg) amino acids. The linkages obtained were also varied, from the imidate-sugar we had an amidine, from the isothiocyanate a thiourea, and from the azide a triazole. As our aim was to find a way to maximize the efficiency and diversity of glycoprotein synthetic pathways by introducing the flexibility to rapidly switch between different strategies, here we present a practical approach which allows the formation of products with different glycosylation sites, linkage types and copy numbers (Table 3.1).

Entry	Protein	Sugar	Glycans added ^a	Conversion % ^b
1	Np276-Hpg61	2.80	2:3	28:72
2	Np276-Hpg61	3.4	1	>95
3	Np276-Hpg61	3.7	1	>95
4	Q β -Hpg16	2.80	1	30
5	Q β -Hpg16	3.4	1	58
6	Q β -Hpg16	3.7	1	>95

^aNumber of sugars added to the protein during each conjugation

^bConversion obtained determined by MS analysis.

Table 3.1. Glycoprotein conversions obtained using the triply divergent approach. For almost all couplings only one unit of GlcNAc was added, obtaining a partial conversion for Q β with both IME and NCS reagents. Only IME coupling on Np276 resulted in multiple glycosylation.

We suggest that methodology can be applied to different (and even more complex) sugars and proteins, also using different types of chemistries, to exploit the advantage it offers in strategic flexibility.

3.7 References

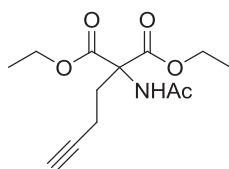
- [1] B. G. Davis, *Chem. Rev.* **2002**, 102, 579–601.
- [2] D. P. Gamblin, E. M. Scanlan, and B. G. Davis, *Chem. Rev.* **2009**, 109, 131–163.
- [3] L. X. Wang and M. N. Amin, *Chem. Biol.* **2014**, 21, 51–66.
- [4] A. Varki, *Glycobiology* **1993**, 3, 97–130.
- [5] B. G. Davis, in *Biotherapeutics: Recent Developments using Chemical and Molecular Biology* (Eds.: by L. H. Jones, A. J. McKnight), Royal Society of Chemistry, Cambridge, **2013**, 130–144.
- [6] J. R. Rasmussen, *Curr. Biol.*, **1992**, 2, 590.
- [7] A. M. Sinclair and S. Elliott, *J Pharm Sci*, **2005**, 94, 1626–1635.
- [8] S. De Munari, T. Schiffner, and B. G. Davis, *Isr. J. Chem.* **2015**, 55, 387–391.
- [9] C. D. Spicer and B. G. Davis, *Nat. Commun.* **2014**, 5, 4740.
- [10] J. M. Chalker, G. J. L. Bernardes, and B. G. Davis, *Acc. Chem. Res.* **2011**, 44, 730–741.
- [11] S. V Sambasivarao, **2013**, 18, 1199–1216.
- [12] P. J. Higgins and H. F. Bunn, *J. Biol. Chem.*, **1981**, 256, 5204–5208.
- [13] N. Floyd, B. Vijayakrishnan, J. R. Koeppe, and B. G. Davis, *Angew. Chem. Int. Ed.* **2009**, 48, 7798–7802.
- [14] M. Lo Conte, S. Staderini, A. Marra, M. Sanchez-Navarro, and B. G. Davis, *Chem. Commun.* **2011**, 47, 11086–88.
- [15] J. M. Chalker, C. S. C. Wood, and B. G. Davis, *J. Am. Chem. Soc.* **2009**, 131, 16346–7.

- [16] C. D. Spicer and B. G. Davis, *Chem. Commun. (Camb)*. **2011**, 47, 1698–700.
- [17] C. D. Spicer, T. Triemer, and B. G. Davis, *J. Am. Chem. Soc.* **2012**, 134, 800–803.
- [18] G. J. L. Bernardes, L. Linderoth, K. J. Doores, O. Boutureira, and B. G. Davis, *Chembiochem* **2011**, 12, 1383–6.
- [19] S. I. van Kasteren, H. B. Kramer, D. P. Gamblin, and B. G. Davis, *Nat. Protoc.*, 2007, 2, 3185–94.
- [20] S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, and B. G. Davis, *Nature*, 2007, 446, 1105–9.
- [21] Y. a Lin, J. M. Chalker, N. Floyd, G. J. L. Bernardes, and B. G. Davis, *J. Am. Chem. Soc.* **2008**, 130, 9642–3.
- [22] M. Vetting and S. Hegde, *Protein Sci.* **2007**, 16, 755–760.
- [23] M. Fernández-González, O. Boutureira, G. J. L. Bernardes, J. M. Chalker, M. A. Young, J. C. Errey, and B. G. Davis, *Chem. Sci.* **2010**, 1, 709.
- [24] M. Kajitani, A. Kato, A. Wada, Y. Inokuchi, and A. Ishihama, *J. Bacteriol.* **1994**, 176, 531–4.
- [25] J. D. Fiedler, C. Higginson, M. L. Hovlid, A. A. Kislukhin, A. Castillejos, F. Manzenrieder, M. G. Campbell, N. R. Voss, C. S. Potter, B. Carragher, and M. G. Finn, *Biomacromolecules* **2012**, 13, 2339–48.
- [26] O. Boutureira, F. D’Hooge, M. Fernández-González, G. J. L. Bernardes, M. Sánchez-Navarro, J. R. Koeppe, and B. G. Davis, *Chem. Commun. (Camb)*. **2010**, 46, 8142–4.
- [27] Y. C. Lee, C. P. Stowell, and M. J. Krantz, *Biochemistry* **1976**, 15, 3956–63.

- [28] Z. Wang, in *Comprehensive Organic Name Reactions and Reagents*, John Wiley & Sons, Inc. **2010**, 3123-3128.
- [29] M. A. Robinson, S. T. Charlton, P. Garnier, X. Wang, S. S. Davis, A. C. Perkins, M. Frier, R. Duncan, T. J. Savage, D. A. Wyatt, S. A. Watson, and B. G. Davis, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101, 14527–32.
- [30] M. K. Patel, B. Vijayakrishnan, J. R. Koeppe, J. M. Chalker, K. J. Doores, and B. G. Davis, *Chem. Commun. (Camb)*. **2010**, 46, 9119–21.
- [31] D. Srimani, M. Feller, Y. Ben-David, and D. Milstein, *Chem. Commun. (Camb)*. **2012**, 1–3.
- [32] M. Breyse and R. Frety, *J. Catal.* **1987**, 154, 144–154.
- [33] N. Fischer, E. D. Goddard-Borger, R. Greiner, T. M. Klapötke, B. W. Skelton, and J. Stierstorfer, *J. Org. Chem.* **2012**, 77, 1760–4.
- [34] L. M. Artner, L. Merkel, N. Bohlke, F. Beceren-Braun, C. Weise, J. Dervede, N. Budisa, and C. P. R. Hackenberger, *Chem. Commun.* **2012**, 48, 522.
- [35] Z. Zhou, C. J. Fahrni, *J. Am. Chem. Soc.* **2004**, 126, 8862-3.
- [36] Y. Nagata, M. M. Burger, *J. Biol. Chem.* **1974**, 3116–3122.

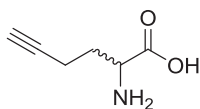
3.8 Supporting information

Diethyl-2-acetamido-2-(but-3-ynyl)malonate **3.1**^[1]



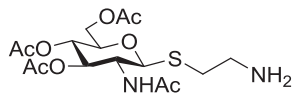
To a solution of NaH (750 mg, 31.25 mmol, 1.1 eq.) in DMF (dry, 30 mL) at 0 °C, under nitrogen, was slowly added a solution of diethyl acetamidomalonate (6.171 g, 28.4 mmol, 1 eq.) in DMF (dry, 20 mL). 4-bromo butyne (2.66 mL, 20 mmol, 0.7 eq.) was added dropwise and the reaction allowed to warm to r.t.. The reaction was stirred at 90 °C for 16 h, then left to cool to r.t. and quenched with a satd. solution of NH₄Cl. The crude mixture was recovered with H₂O (200 mL) and extracted with Et₂O (2 x 150 mL) and EtOAc (2 x 150 mL). The organic layers were combined and washed with brine (200 mL). Finally the crude was dried over MgSO₄, evaporated and purified by flash column chromatography on silica (EtOAc:PE, 1.5:8.5) to afford the title compound **3.1** as a white solid (916.6 mg, 3.4 mmol, 12%). **TLC** R_f = 0.7 (EtOAc, 100%); **¹H NMR** (CDCl₃, 250 MHz) δ 6.84 (s, 1 H, NH), 4.14 (qd, *J* = 7.1, 1.6 Hz, 4 H, 2 x CH₃CH₂O), 2.49 (t, *J* = 7.4 Hz, 2 H, CH₂CH₂CCH), 2.04 (td, *J* = 7.4, 2.6 Hz, 2 H, CH₂CCH), 1.95 (s, 3 H, Ac), 1.86 (t, *J* = 2.7 Hz, 1 H, CH-alkyne), 1.16 (t, *J* = 7.1 Hz, 6 H, 2 x CH₃); **¹³C NMR** (CDCl₃, 63 MHz) δ 169.71 (CONH), 168.05 (2 x CO), 83.00 (CCH-alkyne), 69.42 (CH-alkyne), 66.08 (COCCO), 62.96 (2 x CH₂O), 31.27 (CH₂CH₂CCH), 23.28 (CH₂CCH), 14.25 (2 x CH₃), 13.66 (Ac); **MS** (ESI⁺) *m/z* = 270.30 [M+H]; **IR** (ATR) 3294.94 (CH-alkyne), 1761.18 (COO), 1645.02 (CONH).

D,L-Homopropargylglycine **3.2**^[2,3]



To a solution of **3.1** (916.6 mg, 3.4 mmol, 1 eq.) in MeOH:H₂O (1:1, 16 mL), was added NaOH (157 mg, 3.9 mmol, 1.15 eq.). The reaction was stirred refluxing at 100 °C for 16 h. The reaction was then cooled to r.t., the MeOH evaporated, and the crude recovered with H₂O (100 mL), acidified to pH 2 with HCl (1 M) and extracted with EtOAc (3 x 100 mL). The organic layer was backwashed with brine (150 mL), dried over MgSO₄ and evaporated. The crude was dissolved in HCl (2 M, 60 mL) and refluxed at 100 °C for 16 h. The reaction mixture was left to cool to r.t., loaded on Dowex-H⁺ and the resin was washed with H₂O. The compound was recovered with ammonium hydroxide (5 M) and lyophilised to afford the title compound **3.2** as white solid (186 mg, 1.46 mmol, 44%). **TLC** R_f = 0.0 (EtOAc, 100%); **¹H NMR** (D₂O, 400 MHz) δ 3.83 (dd, *J* = 7.4, 5.4 Hz, 1 H, CH), 2.32-2.43 (m, 2 H, CH₂CCH), 1.97-2.17 (m, 2 H, CH₂CH₂CH); **¹³C NMR** (D₂O, 101 MHz) δ 174.59 (CO), 113.69 (CCH-alkyne), 70.77 (CH-alkyne), 54.01 (CH), 29.42 (CH₂CH₂CCH), 14.36 (CH₂CCH); **MS** (ESI⁺) *m/z* = 128.15 [M+H]; [α]_D²⁰ = +0.0 (c = 1.00, H₂O); **IR** (ATR) 3295.98, 3297.94 (NH₂), 1711.34 (COOH).

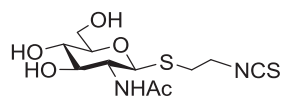
2-N-acetamido-3,4,6-tri-O-acetyl-2-deoxy-1-thio-(S-aminoethyl)- β -D-glucopyranoside **3.3^[4,5]**



To a solution of **2.3** (100 mg, 0.25 mmol, 1 eq.) in EtOH:CHCl₃ (5:1, 60 mL) was added PtO₂ (30 mg) and the mixture was stirred under hydrogen gas at r.t.. After 24 h more PtO₂ (30 mg) was added. After a further 24 h one more portion of PtO₂ (30 mg) was added, along with CHCl₃ (10 mL). After 4 days the reaction mixture was filtered through celite and the solvent was removed under vacuum. The reaction was purified by flash column chromatography on silica (DCM:MeOH, 10:2) to afford the title compound **3.3** as a colourless oil (46 mg, 0.11 mmol, 46%). **TLC** R_f = 0.15 (DCM:MeOH, 10:1); **¹H NMR** (MeOD, 250 MHz) δ 5.22 (t, *J* = 9.7 Hz, 1 H, H-3), 5.02 (t, *J* = 9.7 Hz, 1 H, H-4), 4.75 (d, *J* = 10.5 Hz, 1 H, H-1), 4.22 (qd, *J* = 12.4, 3.6 Hz, 2 H, 2 x H-6), 3.99 (t, *J* = 10.3 Hz, 1 H, H-2), 3.84 (ddd, *J* = 10.0, 4.9, 2.4 Hz, 1 H, H-5), 2.98-2.67 (m, 4 H, SCH₂CH₂), 2.09, 2.04, 2.01, 1.93 (4 x s, 4 x 3 H, 4 x Ac); **¹³C NMR** (MeOD, 101 MHz) δ 171.84, 170.93, 170.40, 169.84 (4 x CO), 84.26 (C-1), 75.54 (C-5), 73.77 (C-3), 68.58 (C-4), 62.04 (C-6), 52.90 (C-2), 41.11 (SCH₂CH₂), 33.22 (SCH₂CH₂), 21.35 (NHAc), 19.30, 19.22, 19.19 (3 x OAc); **MS** (ESI⁺) *m/z* = 407.46 [M+H]; $[\alpha]_D^{20}$ = -95.8 (*c* = 1.0, CHCl₃); **IR** (ATR) 3262.54 (NH₂), 1745.05 (CONH), 1657.36 (COO).

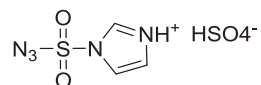
2-N-acetamido-1-thio-(S-isothiocyano-ethyl)-2-deoxy-β-D-glucopyranoside

3.4



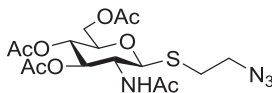
To a solution of **3.3** (50 mg, 0.12 mmol, 1 eq.) in MeOH (dry, 2 mL), was added NaOMe (25% in MeOH, 30 μ L, 0.12 mmol, 1 eq.). The reaction was stirred at r.t. for 1 h, then quenched with DowexH⁺ and the solvent removed under vacuum. The residue was dissolved in NaHCO₃:Na₂CO₃ (1:1, 0.3 M). A solution of thiophosgene (55 μ L, 0.72 mmol, 6 eq.) in THF (2 mL) was added and the mixture stirred at r.t. for 2 h. The solvent was evaporated under nitrogen flow and the product desalted using Biogel P2 resin, eluting with H₂O. The fractions containing desired material were lyophilised to afford the title compound **3.4** as a white solid (43 mg, 0.12 mmol, quant.). **TLC** R_f = 0.0 (EtOAc, 100%); **¹H NMR** (500 MHz, MeOD) δ 4.64 (d, *J* = 10.5 Hz, 1 H, H-1), 3.84-3.79 (m, 2 H, H-2, H-4), 3.71-3.55 (m, 8 H, H-3, SCH₂CH₂, H-6, H-5), 1.90 (s, 3 H, Ac); **¹³C NMR** (126 MHz, MeOD) δ 177.91 (CO), 132.11 (NCS), 83.14 (C-1), 76.42 (C-5), 70.84 (C-3), 64.38 (C-2), 60.87, 60.52 (SCH₂CH₂), 58.35 (C-6), 52.92 (C-4), 20.18 (Ac); **MS** (ESI⁺) *m/z* = 323.41 [M+H]; [α]_D²⁰ = +29.5 (*c* = 1.00, H₂O); **IR** (ATR) 3348.52 (OH), 1694.37 (CO).

Imidazole-1-sulphonyl azide hydrogensulphate **3.5**^[6]



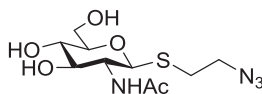
Thionyl chloride (1.6 mL, 22 mmol, 1.1 eq.) was added dropwise to a suspension of sodium azide (1.3 g, 20 mmol, 1 eq.) in ACN (20 mL), at 0 °C. The reaction was stirred overnight at r.t. before imidazole (2.59 g, 38 mmol, 1.9 eq.) was added portion-wise at 0 °C. The reaction was stirred for 3 h at r.t., then diluted with EtOAc (40 mL), washed with H₂O (2 x 40 mL) and satd. solution of NaHCO₃ (2 x 40 mL). The organic layer was dried over MgSO₄ and cooled to 0 °C. A solution of H₂SO₄ (1.1 mL, 20 mmol, 1 eq.) in EtOAc (5 mL) was added dropwise. The reaction was stirred at r.t. for 1 h and the resulting precipitate was removed by filtration, washed with EtOAc, and dried to afford the title compound **3.5** as a white crystalline solid (2.865 g, mmol, 53%). **¹H NMR** (DMSO-d₆, 500 MHz) δ 14.30 (br, 1 H, HSO₄⁻), 10.20 (br, 1 H, NH⁺), 8.79 (s, 1 H, NCHNH), 8.03 (s, 1 H, NCHCHNH), 7.42 (s, 1 H, NCHCHNH); **¹³C NMR** (DMSO-d₆, 126 MHz) δ 138.34 (NCHNH), 130.05 (NCHCHNH), 119.53 (NCHCHNH); **MS** (ESI⁺) m/z = 174.16 [M⁺].

2-N-acetamido-3,4,6-tri-O-acetyl-2-deoxy-1-thio-(S-azidoethyl)- β -D-glucopyranoside **3.6^[5,7]**



To a solution of **3.2** (110.8 mg, 0.27 mmol, 1 eq.) in dry MeOH (3 mL) were added **3.5** (88.7 mg, 0.33 mmol, 1.2 eq.), K₂CO₃ (84.77 mg, 2.25 mmol, 8 eq.) and copper (II) sulphate (0.7 mg, 0.004 mmol, 0.015 eq.). The mixture was stirred under argon at r.t. for 3 h. The reaction was diluted with H₂O (30 mL) and extracted with EtOAc (2 x 30 mL). The organic layer was dried over MgSO₄ and the solvent was removed under vacuum to afford the title compound **3.6** as a white solid (113.3 mg, 0.26 mmol, 97%). **TLC** R_f = 0.9 (DCM:MeOH:NH₄OH, 10:1:0.1); **¹H NMR** (MeOD, 101 MHz) δ 5.11 (t, *J* = 9.7 Hz, 1 H, H-3), 4.90 (t, *J* = 9.7 Hz, 1 H, H-4), 4.69 (d, *J* = 10.5 Hz, 1 H, H-1), 4.14 (dd, *J* = 12.3, 5.1 Hz, 1 H, H-6a), 4.05 (dd, *J* = 12.3, 2.3 Hz, 1 H, H-6b), 3.88 (t, *J* = 10.3 Hz, 1 H, H-2), 3.74 (ddd, *J* = 10.0, 4.9, 2.4 Hz, 1 H, H-5), 3.43 (dd, *J* = 12.9, 6.6 Hz, 2 H, SCH₂CH₂), 2.86, 2.70 (2 x dt, *J* = 13.8, 6.7 Hz, 2 x 1 H, SCH₂CH₂), 1.97, 1.92, 1.89 (3 x s, 3 x 3 H, 3 x OAc), 1.81 (s, 3 H, NHAc); **¹³C NMR** (MeOD, 400 MHz) δ 171.92, 170.95, 170.44, 169.88 (4 x CO), 84.01 (C-1), 75.51 (C-5), 73.68 (C-3), 68.65 (C-4), 62.07 (C-6), 52.88 (C-2), 51.18 (SCH₂CH₂), 29.47 (SCH₂CH₂), 21.27 (NHAc), 19.28, 19.20, 19.17 (3 x OAc); **MS** (EI⁺) *m/z* = 432.45 [M⁺]; $[\alpha]_D^{20}$ = -93.9 (c = 1.0, CHCl₃); **IR** (ATR) 1742.65 (COO), 1662.08 (CONH).

2-N-acetamido-2-deoxy-1-thio-(S-azidoethyl)- β -D-glucopyranoside **3.7**^[5]



To a solution of **3.6** (100 mg, 0.23 mmol, 1 eq.) in MeOH (dry, 2.5 mL) was added NaOMe (25% in MeOH, 53 μ L, 0.23 mmol, 1 eq). The reaction was stirred at r.t. for 1 h before being evaporated under vacuum, to afford the title compound **3.7** as a white solid (71 mg, 0.23 mmol, quant.). **TLC** R_f = 0.1 (DCM:MeOH:NH₄OH, 10:1:0.1); **¹H NMR** (500 MHz, MeOD) δ 4.67 (d, J = 10.5 Hz, 1 H, H-1), 3.87-3.81 (m, 2 H, H-2, H-4), 3.74-3.52 (m, 8 H, H-3, SCH₂CH₂, H-6, H-5), 1.92 (s, 3 H, Ac); **¹³C NMR** (126 MHz, MeOD) δ 177.82 (CO), 83.17 (C-1), 77.56 (C-5), 70.11 (C-3), 66.44 (C-2), 60.95, 60.67 (SCH₂CH₂), 60.11 (C-6), 52.76 (C-4), 20.25 (Ac); **MS** (ESI⁻) m/z = 305.33 [M-H]; $[\alpha]_D^{20}$ = -97.7 (c = 1.0, CHCl₃); **IR** (ATR) 3378.45 (OH), 1667.51 (CO).

Np276-M61Hpg expression

Plasmid carrying the Np276 G2P M21I Single M61 gene was transformed into *E. coli* B834 (DE3) competent cells, plated on LB-agar supplemented with kanamycin (50 μ g/mL) and incubated at 37 °C for 18 h. A single colony from the plate was used to inoculate SelenoMet media (20 mL) supplemented with SelenoMet nutrient mix, L-methionine (20 mg/L) and kanamycin (50 μ g/mL) and incubated for 16 h at 37 °C. Starter cultures were used to inoculate SelenoMet media (800 mL), supplemented with SelenoMet nutrient mix, L-methionine (20

mg/L) and kanamycin (50 µg/mL), the cells incubated at 37 °C, and the optical density (OD₆₀₀) monitored until a value of 0.6-0.8 was obtained (~3-4 h). The cells were then pelleted by centrifugation (8 min, JLA-9.1000 rotor, 8000 rpm). The pellets were washed three times with SelenoMet base media (3 × 400 mL) containing kanamycin (50 µg/mL) and then transferred to pre-warmed SelenoMet media (800 mL) containing kanamycin (50 µg/mL) and supplemented with SelenoMet nutrient mix as well as the desired unnatural amino acid (DL-Hpg, 300-500 mg/L). The culture was incubated for 30 min at 37 °C. Expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG, 1 mM) and was continued for 16 h at 30 °C with shaking (220 rpm). The cells were harvested by centrifugation (8 min, 8000 rpm) to give pellets, which were collected in a 50 mL falcon tube and resuspended in binding buffer (Tris-HCl 20 mM, imidazole 15 mM, NaCl 0.5 M, 1% glycerol, pH 7.8, 20 mL). The cell suspension was treated with DNase (10 µL), and proteinase inhibitor tablet (1/4 tablet), and rocked at 4 °C for 3 h. The cells were then lysed by sonication (Soniprep at 10 micron amplitude, 30 sec on with 30 sec rest, 5 cycles, on ice). The cell debris was removed from the lysis mixture by centrifugation (30 min, JA 25.5 rotor, 20,000 rpm). The protein was purified by incubation of the cleared cell lysate with HisBind Ni-NTA resin (3 mL per litre of culture, Novagen) for 1 h at 4 °C. The resin was packed into a glass econo-column and unbound protein was washed out with binding buffer (A) (Tris-HCl 20 mM, imidazole 15 mM, NaCl 0.5 M, 1% glycerol, pH 7.8) (15-20 CV, ~50 mL). The bound protein was eluted from the resin by elution buffer (B) (Tris-HCl 20 mM, imidazole 500 mM, NaCl 0.5 M, 5% glycerol, pH 7.8) using a step gradient (20% B (5 mL), 40% B (5 mL), 60% B (5 mL), 80%

B (5 mL), 100% B (4 x 5 mL)). All eluted proteins were collected in 5 mL fractions, which were analysed by SDS-PAGE (Fig.S1).

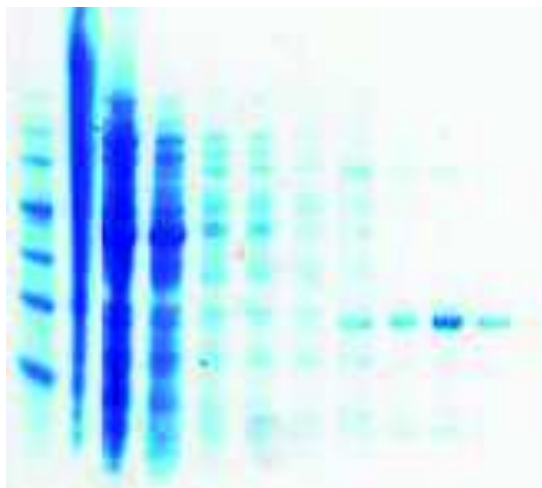


Figure S1. SDS-PAGE gel of Np276-M61Hpg purification. From the left: ladder (first) and elution fractions. Single spots are considered pure protein.

Fractions containing pure Np276 protein were combined and subjected to ammonium sulphate precipitation (85% saturation) for 16 h at 4 °C. The precipitated proteins were pelleted by centrifugation (40 min, JA 25.5 rotor, 20,000 rpm, 4 °C), desalted by overnight dialysis against sodium phosphate buffer (PB) (50 mM, pH 7.5) then analysed by LC-MS.

Protein concentration for Np276-M61Hpg was determined by BCA assay (0.55 mg/mL). The protein was then analysed by LC-MS (Calc: 20855, Found: 20853, Fig. S2).

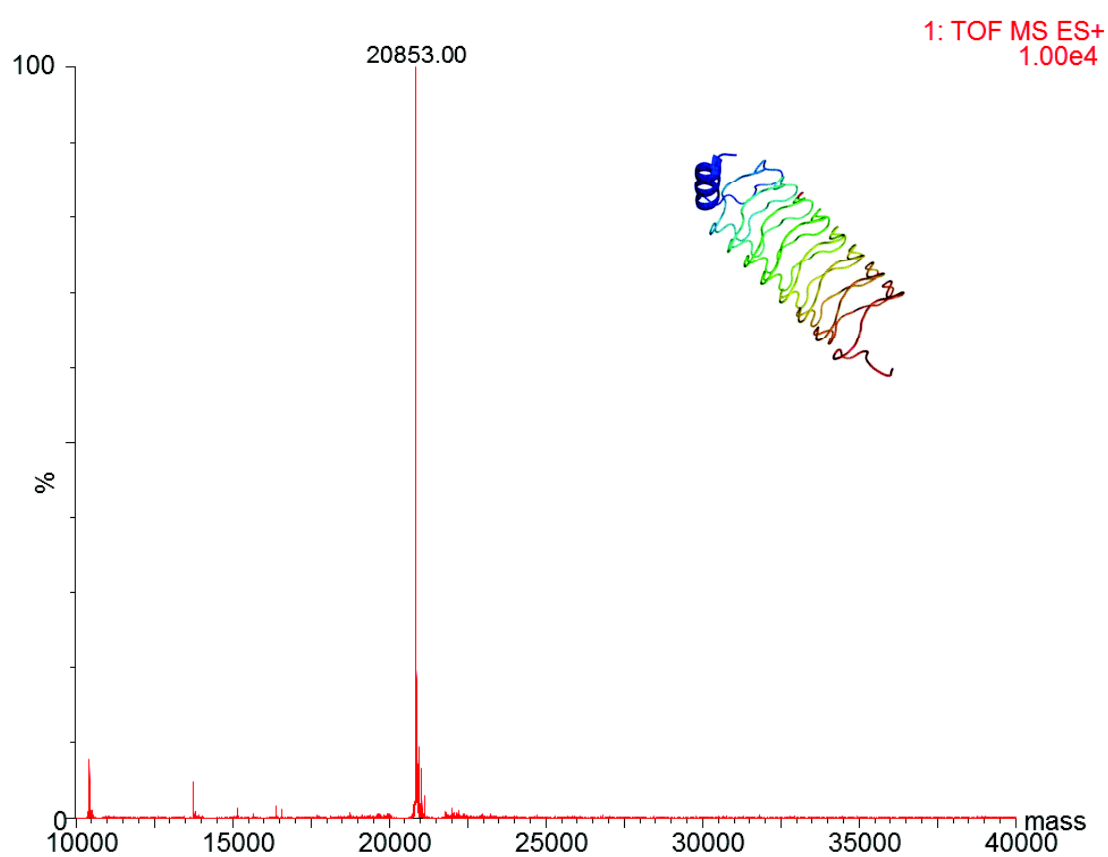
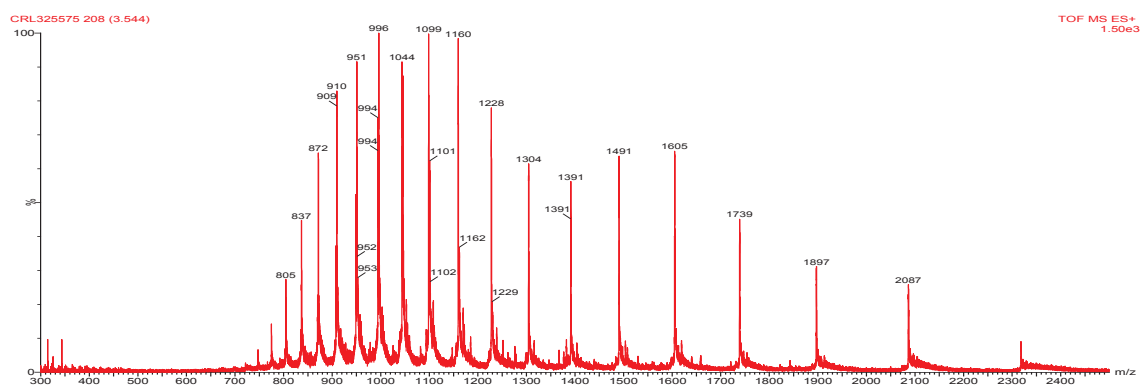


Figure S2. ESI-MS of pure Np276-M61Hpg. Above is the ion series before deconvolution.

Q β -M16Hpg expression

An *E. coli* B834 (DE3) glycerol stock of Q β carrying the K16M gene and p75M vector was used to inoculate SelenoMet media (20 mL) containing ampicillin (100 μ g/mL) and supplemented with SelenoMet nutrient mix and L-methionine (40 μ g/mL). The overnight culture was used to inoculate SelenoMet media (800 mL) containing ampicillin (100 μ g/mL) and supplemented with SelenoMet nutrient mix and L-methionine (40 μ g/mL). The culture was incubated (37 °C, 180 rpm) and grown to OD₆₀₀ between 0.6 and 0.8. The cells were then pelleted by centrifugation (7000 rpm, 8 min). The pellet was washed with SelenoMet media (3 $\times$ 400 mL) containing ampicillin (100 μ g/mL). Cells were then transferred to pre-warmed (37 °C) SelenoMet media (800 mL) containing ampicillin (100 μ g/mL) and supplemented with SelenoMet nutrient mix as well as the desired unnatural amino acid (DL-Hpg, 300-500 mg/L) in place of methionine. The culture was incubated for 30 min at 37 °C. Expression was induced by addition of IPTG (1 mM) and was continued for 6 h at 37 °C with shaking (180 rpm). The cells were harvested by centrifugation (7000 rpm, 8 min) and were resuspended in TBS buffer (Tris-HCl 20 mM, NaCl 150 mM, EDTA 2 mM, pH 7.5, 20 mL). The cells were sonicated (Soniprep at 10 micron amplitude, 30 sec with 1 min rest, 5 times, on ice) and the cell debris removed by centrifugation (20,000 rpm, 30 min). The clear lysate was added to an equal volume of *n*-BuOH:CHCl₃ (1:1) and mixed by inversion. The layers were separated by centrifugation (3,750 rpm, swing-out, 30 min). The aqueous layer was isolated and the virus-like particles were precipitated by addition of PEG8000 (10% w/v) and NaCl (100 mM), dissolved by vortexing and the suspension rocked at 4 °C for 30 min. The precipitated Q β

particles were collected by centrifugation (3750 rpm, swing-out, 30 min) and re-suspended in TBS buffer (10 mL). The extraction step was repeated a second time on the re-suspended Q β . The resulting aqueous layer was isolated and the protein purified by sucrose gradient (5-40%) using an ultracentrifuge (37,600 rpm, 3 h, Beckman Coulter XL-80, SW-40-Ti rotor). Fractions were analysed by native gel (Fig.S3) and pooled together as appropriate.

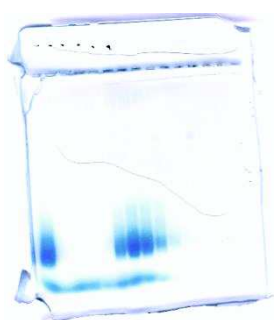


Figure S3. Native gel of Q β -M16Hpg purification. From the left: reference (first) and elution fractions. Single spots are considered pure protein.

The protein was dialysed overnight into TBS buffer (50 mM pH 7.5) and DLS (dynamic light scattering) analysis (Fig. S4) was recorded on a Viscotek 801 DLS.

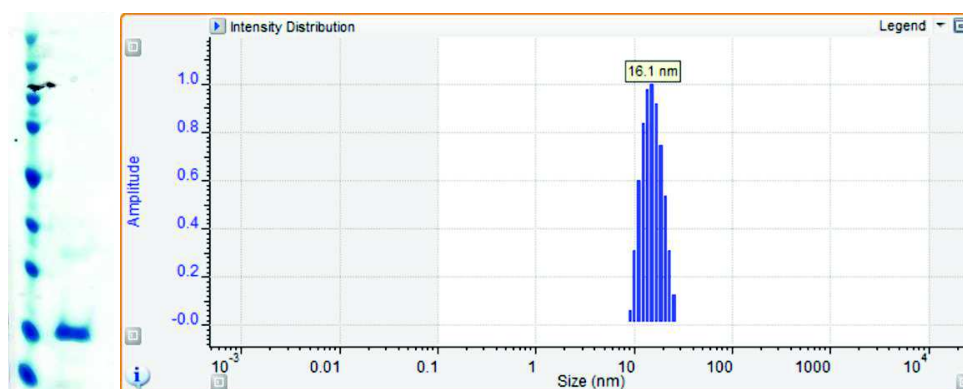


Figure S4. SDS-PAGE gel of pure Q β -M16Hpg (left, in denaturing conditions) and DLS showing a narrow distribution of the particle size as 180-copy multimeric protein.

Protein concentration for Q β -M16Hpg was determined by BCA assay (1.83 mg/mL). The protein was then analysed by LC-MS (for each monomer, Calc: 14104, Found: 14104, Fig.S5).

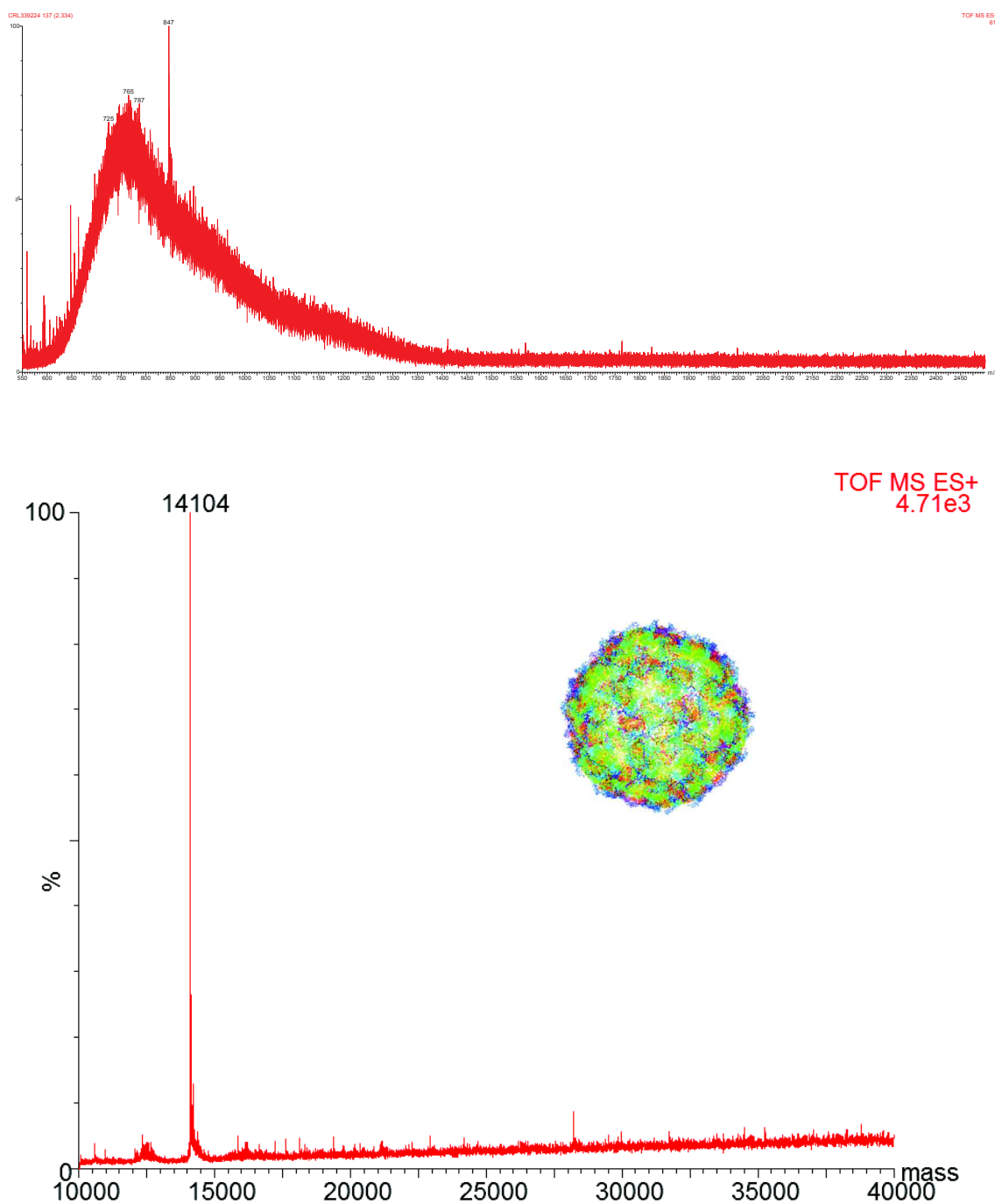
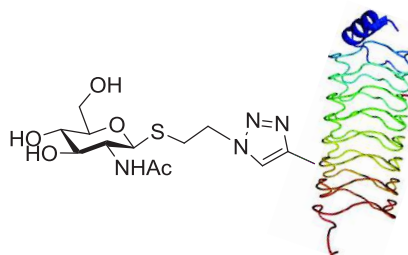


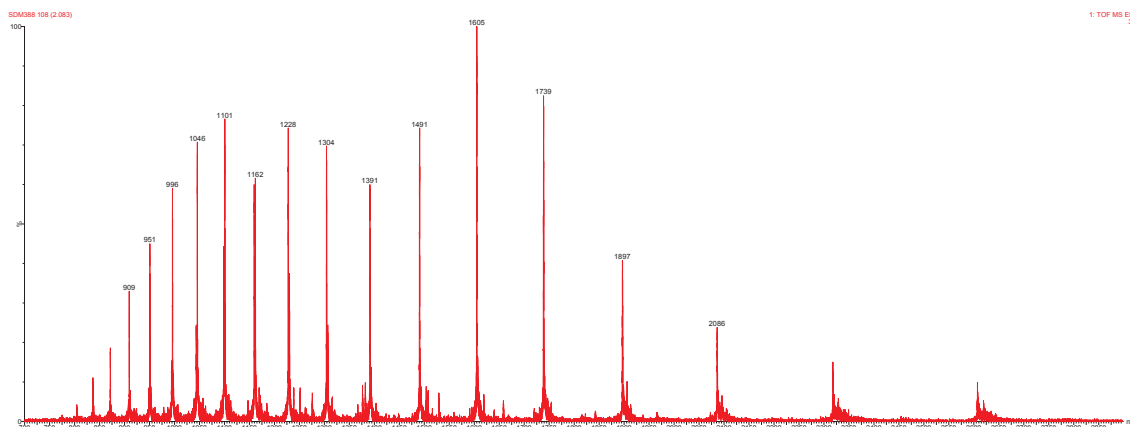
Figure S5. ESI-MS of pure Q β -M16Hpg. Above is the ion series before deconvolution.

CuAAC on Np276-M61Hpg



To Np276-M61Hpg (0.55 mg/mL in PB, 50 mM, pH 8, 200 μ L) was added **3.7** (1.6 mg, 5 μ mol, 1000 eq.). A pre-mixed solution of CuBr (10 mg/mL in ACN, 22.8 μ L) and TTMA (100 mg/mL in ACN, 17.2 μ L) was added to the reaction mixture every 30 min, for a total of 3 additions. The reaction was shaken at r.t. for 30 min and briefly centrifuged to separate the copper salts. The crude mixture was desalted by PD Minitrapp-G25.

The protein was then analysed by LC-MS, which showed full conversion to the desired product (Calc: 21162, Found: 21160, Fig.S6).



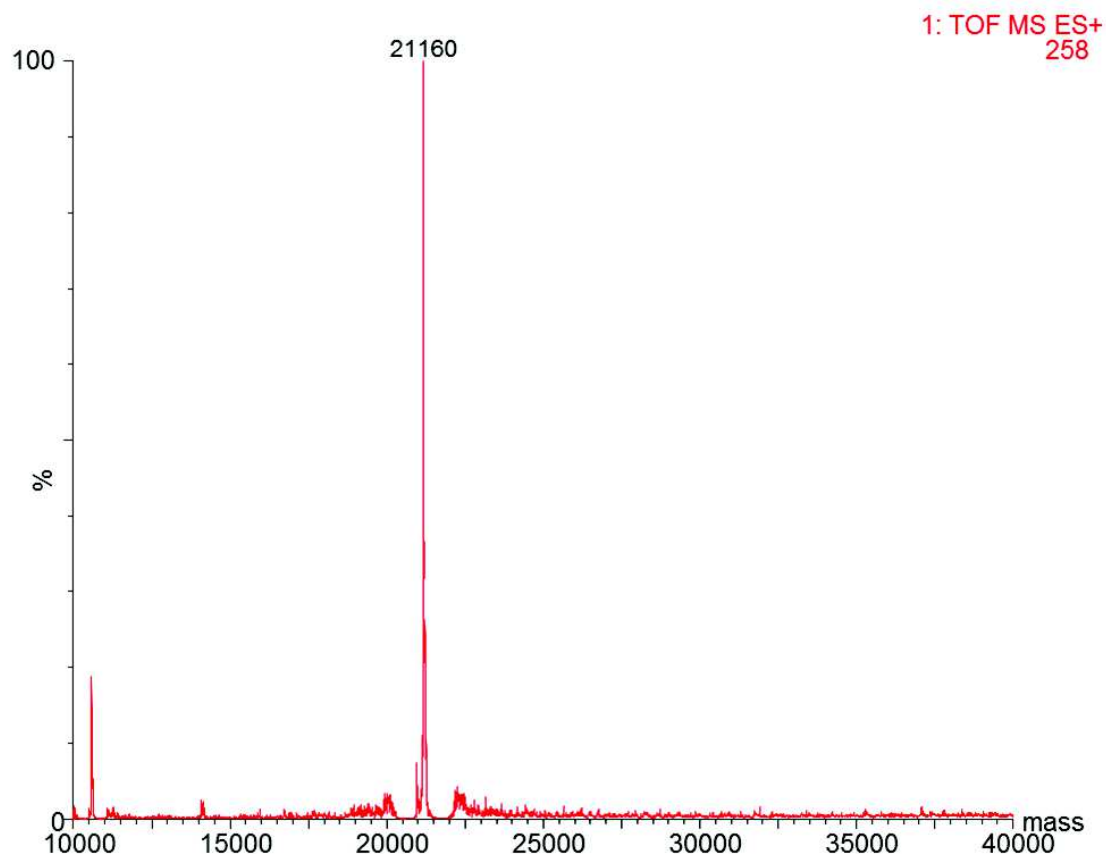
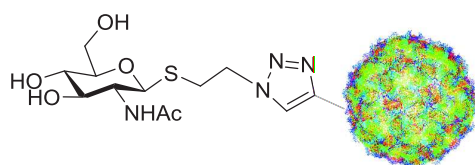


Figure S6. ESI-MS of pure Np276-M61Hpg after CuAAC reaction. Above is the ion series before deconvolution.

CuAAC on Q β -M16Hpg



To Q β -M16Hpg (1.83 mg/mL, in PB, 50 mM, pH 8, 50 μ L) was added **3.7** (2 mg, 6 μ mol, 1000 eq.). A pre-mixed solution of CuBr (10 mg/mL in ACN, 22.8 μ L) and TTMA (100 mg/mL in ACN, 17.2 μ L) was added to the reaction mixture every 30 min, for a total of 5 additions. The reaction was shaken at r.t. overnight and briefly

centrifuged to separate the copper salts. The crude was then desalted by PD Minitrap-G25.

The protein was then analysed by LC-MS, which showed full conversion to the desired product (for each monomer, Calc: 14410, Found: 14410, Fig. S7).

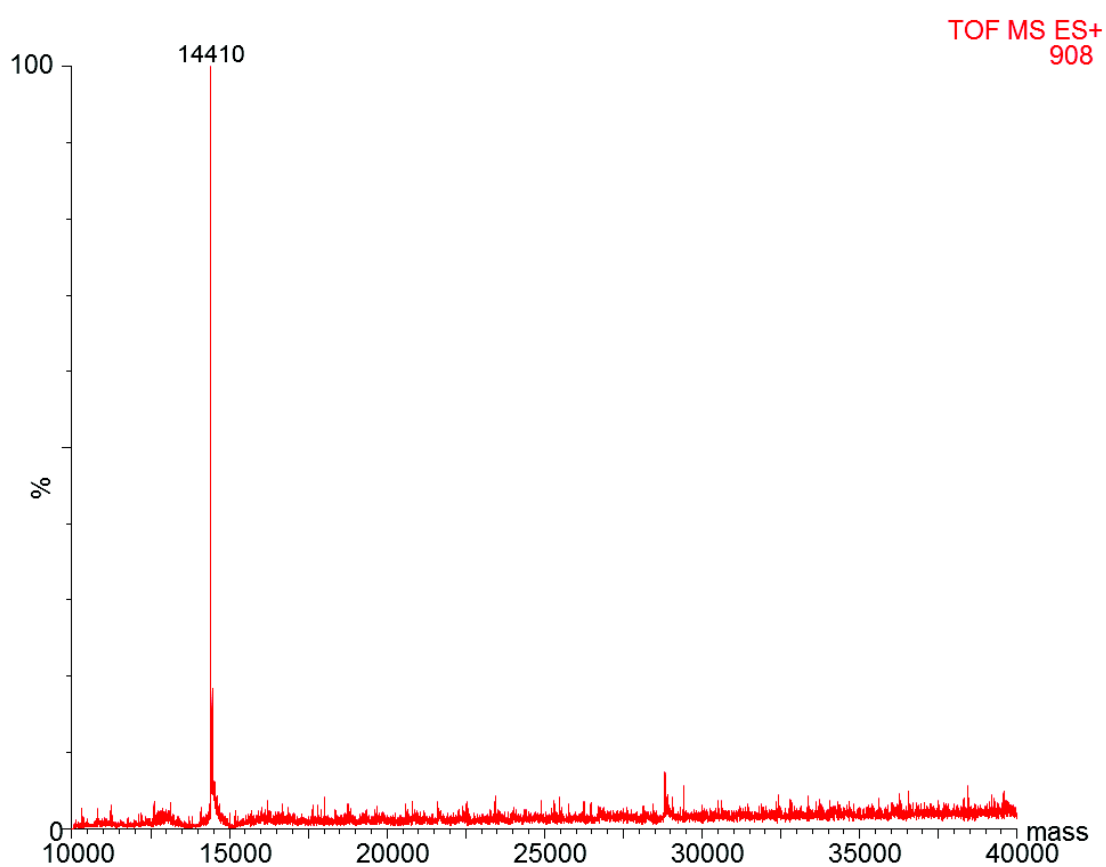
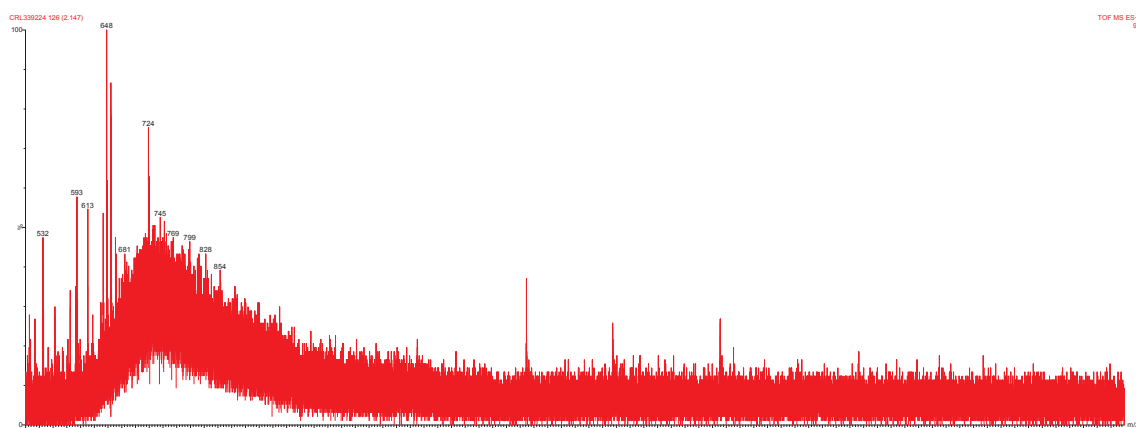
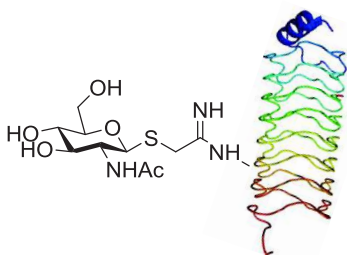


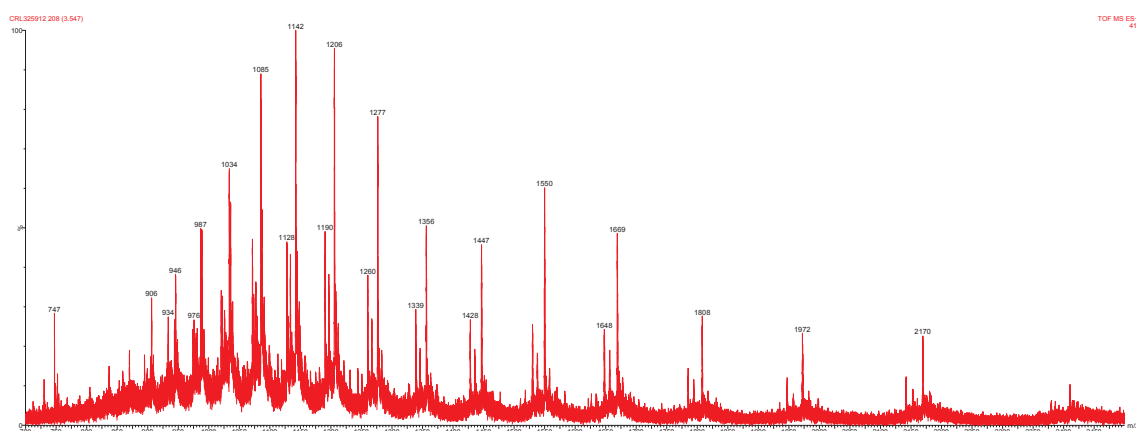
Figure S7. ESI-MS of pure Q β -M16Hpg after CuAAC. Above is the ion series before deconvolution.

IME on Np276-M61Hpg



To Np276-M61Hpg (0.55 mg/mL in PB, 50 mM, pH 8, 200 μ L) was added **2.80** (1.5 mg, 5 μ mol, 1000 eq.) dissolved in TAPS buffer (0.6 M, pH 9). The reaction was shaken at r.t. for 2 h, then desalted by PD Minitrap-G25.

The protein was then analysed by LC-MS, which showed full conversion to the desired product (Calc: 21407-di, 21683-tri, Found: 21409, 27%, and 21685, 73%, Fig. S8).



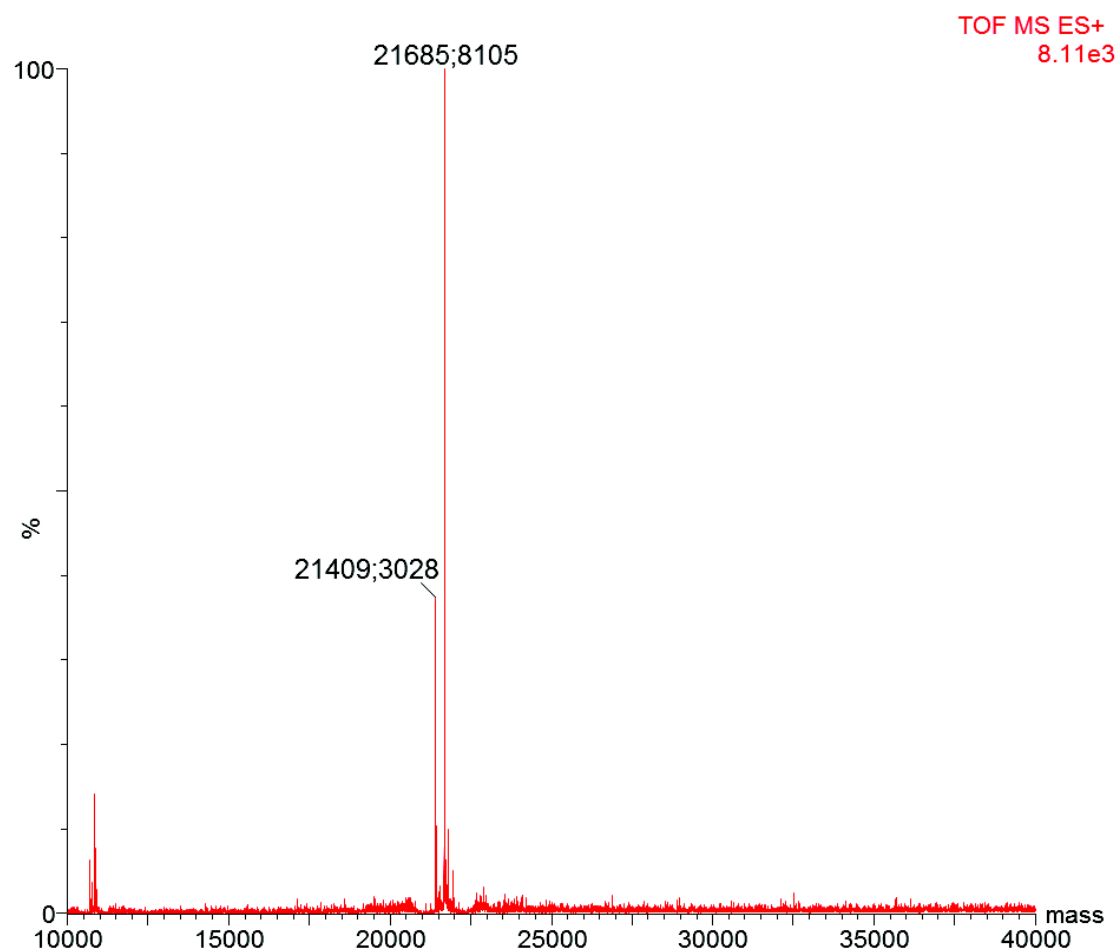
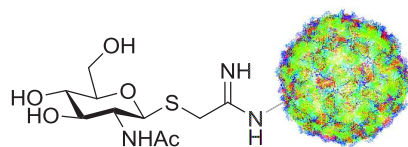


Figure S8. ESI-MS of pure Np276-M61Hpg after IME coupling. Above is the ion series before deconvolution.

IME on Q β -M16Hpg



To Q β -M16Hpg (1.83 mg/mL, in PB, 50 mM, pH 8, 50 μ L) was added **2.80** (1.8 mg, 6 μ mol, 1000 eq.). The reaction was shaken at r.t. for 1 h, then desalted by PD-10.

The protein was then analysed by LC-MS, which showed partial conversion to the desired product (for each monomer, Calc: 14373-mono, Found: 14104, 74%, and 14381, 26%, Fig. S9).

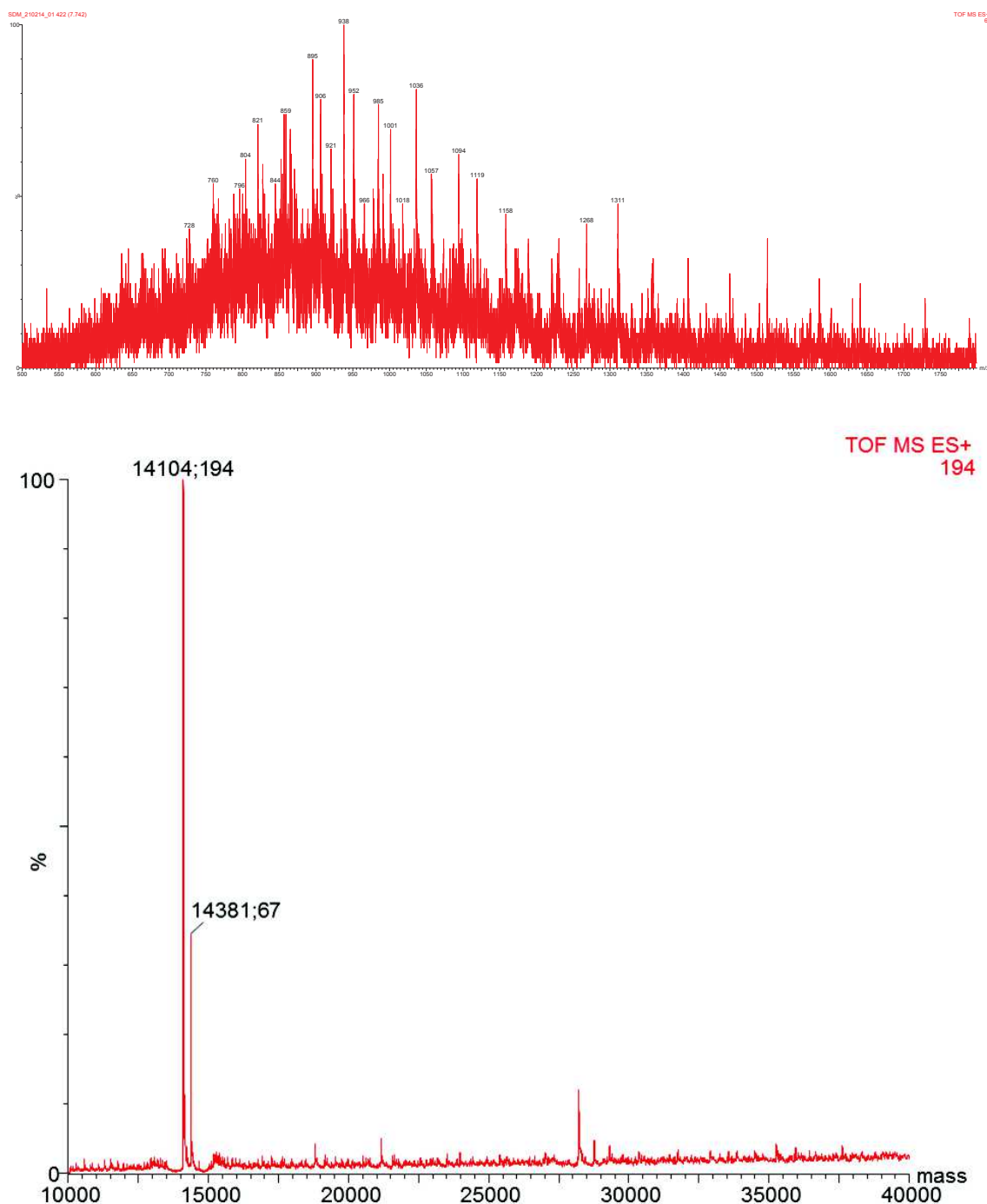
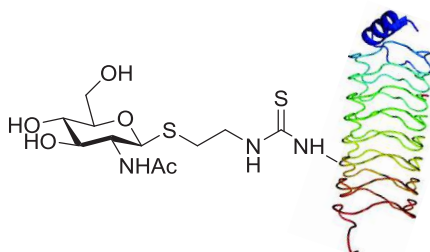


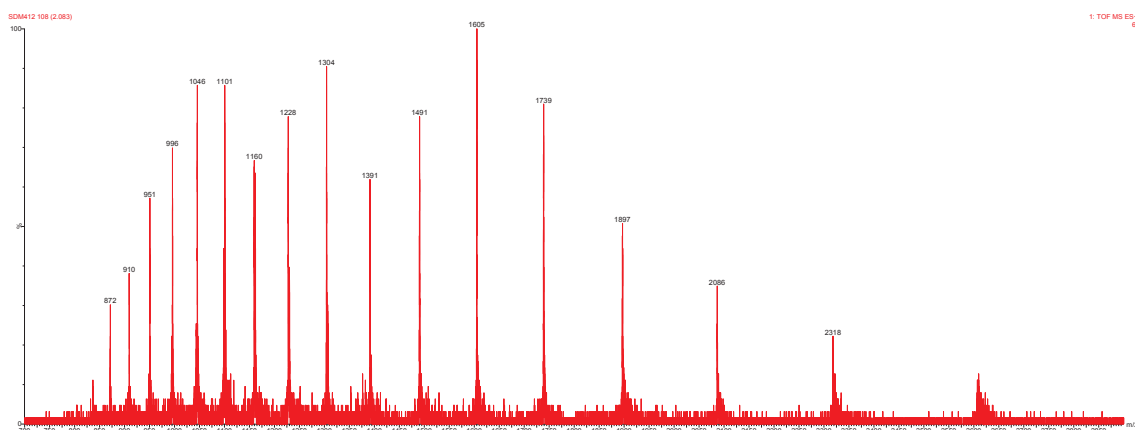
Figure S9. ESI-MS of pure Qβ-M16Hpg after IMC coupling. Above is the ion series before deconvolution.

NCS on Np276-M61Hpg



To Np276-M61Hpg (0.55 mg/mL in PB, 50 mM, pH 8, 200 μ L) was added **3.4** (8 mg, 25 μ mol, 5000 eq.) in NaHCO₃ (0.3 M, 200 μ L). The reaction was shaken at r.t. overnight, then desalted by PD Minitrapp-G25.

The protein was then analysed by LC-MS, which showed full conversion to the desired product (Calc: 21177, Found: 21174, Fig. S10).



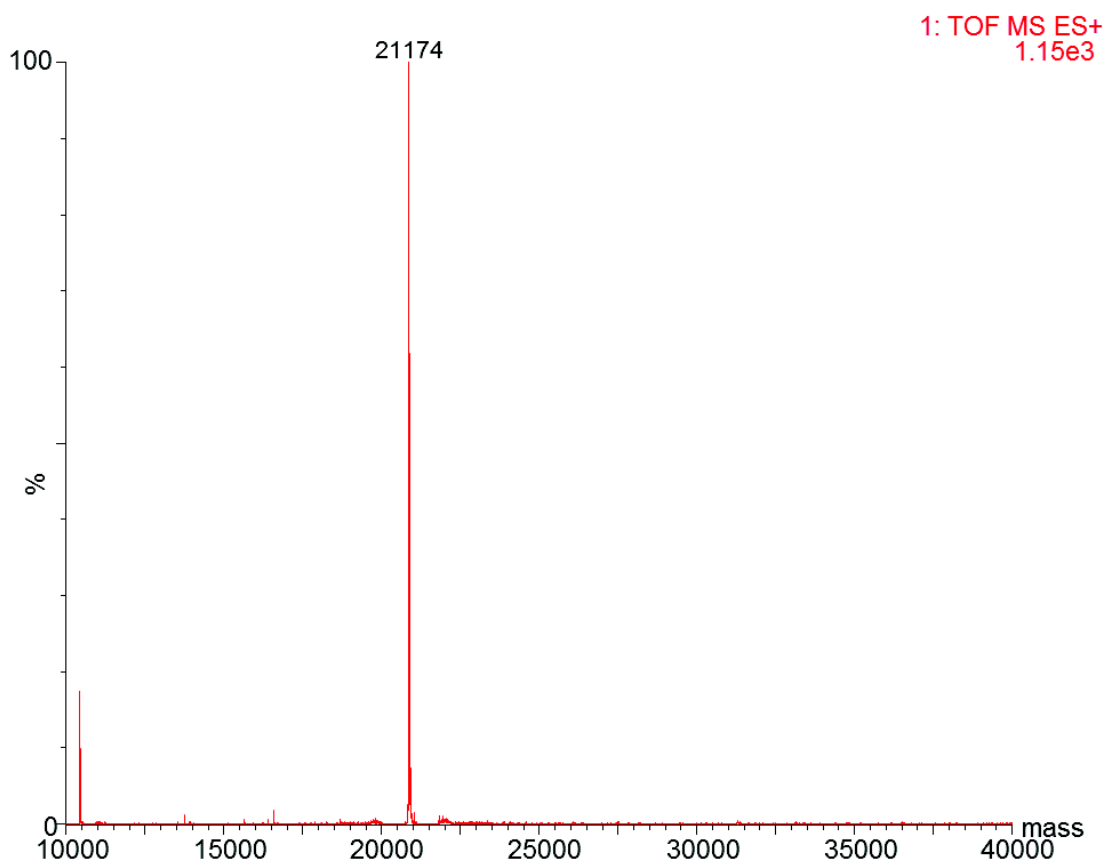
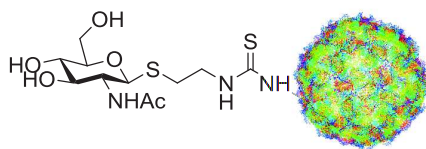


Figure S10. ESI-MS of pure Np276-M61Hpg after NCS coupling. Above is the ion series before deconvolution.

NCS on Q β -M16Hpg



To Q β -M16Hpg (1.83 mg/mL, in PB, 50 mM, pH 8, 50 μ L) was added **3.4** (9.7 mg, 30 μ mol, 1000 eq.) in NaHCO₃ (0.3 M, 200 μ L). The reaction was shaken at r.t. overnight, then desalted by PD Minitrap-G25.

The protein was then analysed by LC-MS, which showed partial conversion to the desired product (for each monomer, Calc: 14425, Found: 14104, 40%, and 14425, 60%, Fig. S11).

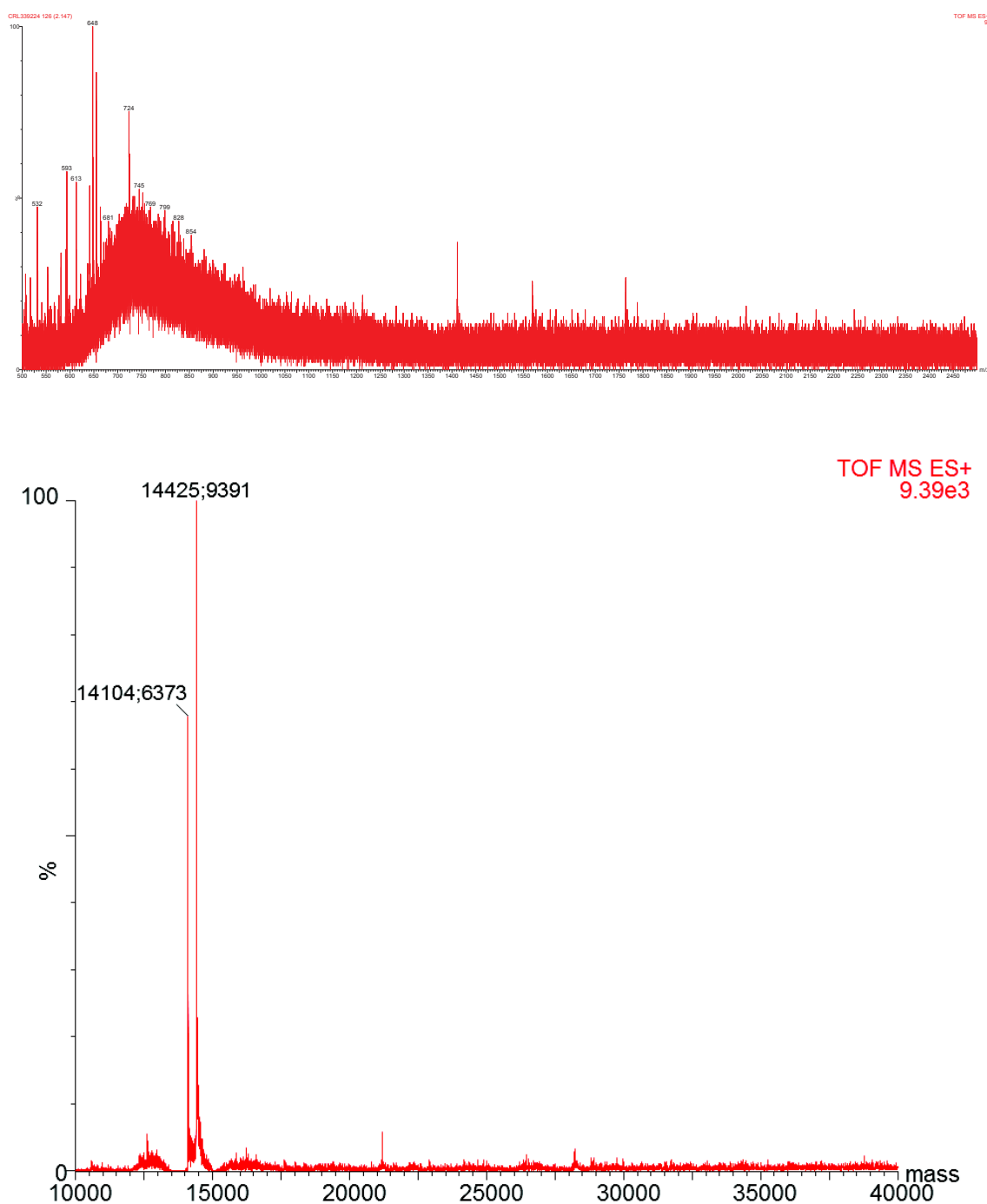


Figure S11. ESI-MS of pure Q β -M16Hpg after NCS coupling. Above is the ion series before deconvolution.

Binding assay

~4,000 response units (RUs) of streptavidin were immobilized on two channels of a CM5 chip using the amine coupling kit according to the manufacturer's instructions (GE Healthcare). Briefly, surfaces were cleaned with NaOH (50 mM) for 30 seconds and activated by passing a mixture of EDC hydrochloride (final concentration 200 mM) and NHS (final concentration 50 mM) over the surface for 7 min. Streptavidin was immobilised on the surface for 7 min at 60 µg/mL in sodium acetate (10 mM, pH 4.5, GE Healthcare). Remaining binding sites were blocked with ethanolamine (1 M) for 7 min. ~2,500 RUs of biotinylated wheat germ agglutinin (WGA; Vector Laboratories Ltd, Peterborough, United Kingdom) were captured onto one flow cell by passing a 20 µL/min solution over the surface for 7 min. Remaining binding sites were blocked by floating biotin (100 mM, Avidity, Aurora, USA) over both the WGA and the streptavidin control surface. Binding experiments were performed at 25 °C at a flow rate of 20 µL/min. Samples were passed over the surfaces at 20 µg/mL for 300 seconds followed by a 300 second dissociation period and surfaces were regenerated by two sequential 30 second injections of glycine (10 mM, pH 2.0, GE Healthcare). Data were analysed by double referencing using a blank control flow cell as well as a buffer injection in the BIAevaluation software (version 4.0.1; Biacore) and curves were plotted in Prism v6.04 (GraphPad Software, La Jolla, USA).

References

- [1] K. L. Kiick and D. A. Tirrell, *Tetrahedron* **2000**, 56, 9487–9493.
- [2] J. C. M. Van Hest, K. L. Kiick, and D. A. Tirrell, *J. Am. Chem. Soc.* **2000**, 122, 1282–1288.
- [3] Y. Aoyagi, T. Sugahara, *Phytochemistry* **1985**, 24, 1835–1836.
- [4] Y. Inaba, T. Fujimoto, H. Ono, M. Obata, S. Yano, and Y. Mikata, *Carbohydr. Res.* **2008**, 343, 941–50.
- [5] S. De Munari, T. Schiffner, and B. G. Davis, *Isr. J. Chem.* **2015**, 55, 387–391.
- [6] N. Fischer, E. D. Goddard-Borger, R. Greiner, T. M. Klapötke, B. W. Skelton, and J. Stierstorfer, *J. Org. Chem.* **2012**, 77, 1760–4.
- [7] E. D. Goddard-Borger and R. V. Stick, *Org. Lett.* **2007**, 9, 3797–800.

Iron nanoparticles as multimeric glycosylated scaffolds

4.1 Introduction

4.1.1 Molecular imaging techniques: PET and MRI

Different techniques are available nowadays for *in vivo* imaging of molecules and biological events.^[1,2] A popular technique is positron emission tomography (PET) which has high resolution and produces a three-dimensional image of functional processes in the body. This imaging tool had its breakthrough only in the 1950s, and it is now used both in research and for medical as well as clinical applications. ^[3] PET is based on the detection of gamma rays emitted by a radionuclide (tracer), introduced into the body with a targeting moiety. Radionuclides selected for this application usually have short half-lives (e.g. ^{11}C with $t_{1/2}$ of 20 min, ^{18}F with $t_{1/2}$ of 110 min). ^[4] In fact, even though this is not an invasive technique it does involve exposure to ionising radiations, which needs to be minimised. Also, radioactive samples need to be prepared and handled by trained personnel in dedicated facilities. Due to the short half-life of the isotopes used, they are often prepared at the same site where they will be administered to minimise their decay before the analysis itself will be carried out. One other disadvantage of this technique, not directly related to the use of radionuclides, is the operating costs.

Despite all these considerable disadvantages, the quality, resolution and potential of this technique have made PET applications in cardiology, neurology and oncology flourish wildly. ^[5-7] To date, the most used PET agent is probably [¹⁸F]-fluorodeoxyglucose ([¹⁸F]-FDG), ^[8-10] an analogue of glucose where the C-2 hydroxyl group has been replaced with the radioactive fluorine.

Among other popular imaging tools, magnetic resonance imaging (MRI) is a non-invasive technique that can penetrate deep tissues. ^[11] This technique was first introduced in the scientific panorama by Paul C. Lauterbur in 1973^[12] and has been made practical for medical investigations by Peter Mansfield,^[13] earning them both the 2003 Nobel Prize in Physiology or Medicine. MRI is based on the measurement of the relaxation rate (T₁ and T₂) of molecules containing hydrogen (such as water, proteins and lipids) when perturbed by a magnetic pulse. Similarly to the mechanism by which organic chemists record an NMR, hydrogen atoms absorb and re-emit electromagnetic radiation, to form a detailed image of internal body tissues. The contrast between different tissues is determined by the rate at which excited atoms return to the equilibrium state, which strongly depends on the environment around the atoms. With the aid of contrast agents,^[14,15] paramagnetic substances shortening the T₁ rate or speeding up the T₂ rate (such as gadolinium, iron and manganese), the sensitivity of MRI can be greatly increased. For example, iron oxide coated nanoparticles present strong T₂-relaxing effects due to the paramagnetic metal core. Under a strong magnetic field, each iron molecule will align, inducing a huge net magnetic moment and a local magnetic field that will decrease the T₂ relaxation rate of the water molecules around them and darken the region where the particles have accumulated. A variety of targeted-MRI contrast agents for molecular imaging have been

made,^[16,17] such as protein-targeted Gd contrasts^[18] and small molecule-Gd derivatives for detection of atherosclerotic plaques;^[19] some of them are already used in clinical diagnosis.^[20,21] As no ionizing radiations are used in MRI, it is preferred over X-ray computed tomography (CT) when both approaches can give comparable diagnostic information. Unfortunately, MRI is more expensive, has limited sensitivity compared to other imaging techniques, takes longer than CT and needs strong magnetic fields, so it cannot be used on patients with metal implants.

4.1.2 Glyconanoparticles as multivalent scaffolds

Iron-oxide nanoparticles (NPs) are commercially available with a polymer coating (e.g. dextran, Fig. 4.1) that increases their stability, makes them water soluble, and due to the functional groups that the polymer exposes on the surface, enables their functionalisation turning them into a multivalent scaffold.^[22]

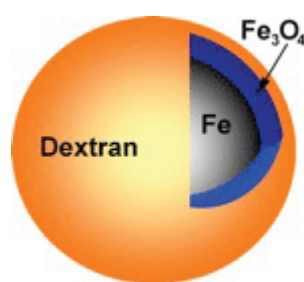


Fig.4.1. Example of iron coated NPs

Functionalised iron NPs have been tested *in vivo* for multiple targeting purposes^[23,24] and their toxicity has been claimed to be lower than gadolinium based contrast agents.^[25] We decided to use amino-functionalised polystyrene coated iron beads (NanoLink™, ~800 nm, 40% iron content, 41.26 nmol/mg of

NH₂) as a multivalent MRI-active platform for our library of saccharides (see Chapter 2).

Once functionalised, these glycosylated contrast agents display highly packed carbohydrates, taking advantage of the cluster glycoside effect^[26] and therefore compensate for the low sensitivity and low avidity for carbohydrate-binding proteins that is typical of constructs with low sugar copy number.^[27] We envisage that these GNPs will provide a versatile tool to explore the *in vivo* distribution of the synthesised sugars, and contribute to the vision of mapping their biological interactions.

4.1.3 Dual MRI-PET probes

Multiple imaging techniques are often used simultaneously,^[28,29] to obtain clear images and accurate diagnosis: the only examples of dual-modal NPs reported so far are fluorescent magnetic NPs.^[30] Here we decided to exploit the idea of a dual MRI-PET probe^[31] by labelling the functionalised iron NPs with ¹⁸F.

Previously, NPs decorated with sLe^x were tested *in vivo* against brain inflammation^[32] by MRI. sLe^x is the sialylated derivative of Le^x, a member of a family of blood group antigens known as the Lewis saccharides.^[33] The role of the four Lewis saccharides (Le^x, Le^y, Le^a, Le^b) during inflammation processes is well-known,^[34] but the distribution of their receptors and the extent by which each of them is involved in the inflammation response is still uncertain. Preliminary data revealed a striking difference in the behaviour of Le^a GNPs with respect to the other Lewis-NPs in a new set of experiments:^[35] the left side of a rat's brain was inoculated with interleukin 1 β (IL-1 β) and the inflammation allowed to

develop before injecting the contrast via tail vein. Le^a-GNPs accumulated in the inflammation site showing a strong binding to the activated endothelium while no signal was detected for the other Lewis-GNPs (Fig. 4.2), or in the right hemisphere (control).

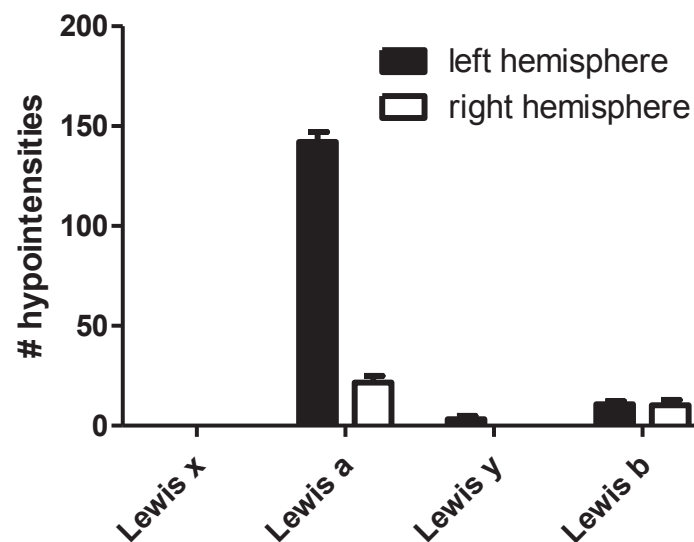
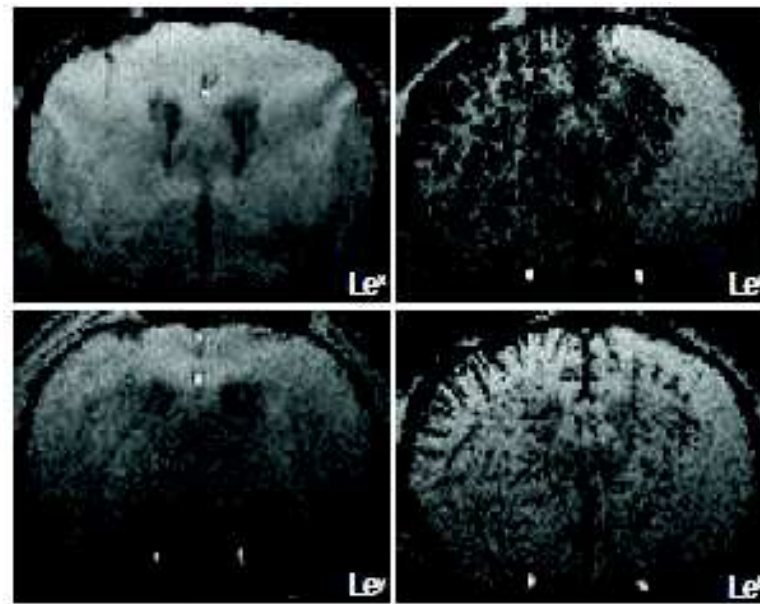


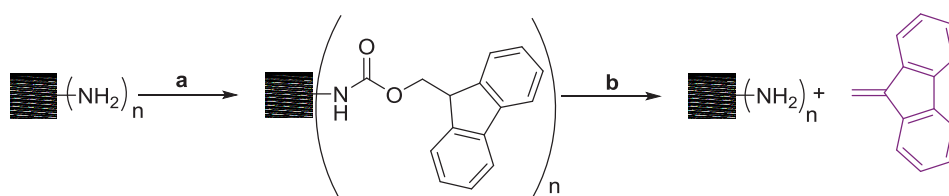
Fig.4.2. a) Minimum intensity projections from in vivo MRI of Lewis-GNP after intracerebral injection of IL-1 β in the rat left brain hemisphere; b) Quantification of the black pixels in the left (IL-1 β) and right (control) brain hemispheres from minimum intensity projections (mean of independent experiments, bars: SEM).

After coupling, GNPs were separated from the reaction mixture using a magnet and repeatedly washed to remove all the unreacted sugars. The magnetic behaviour of the beads facilitated this task as holding the magnet to the side walls of the reaction vial was enough to attract and retain the NPs while the solution was easily removed from the vial by suction with a pipette.

4.3 Characterisation of glyconanoparticles

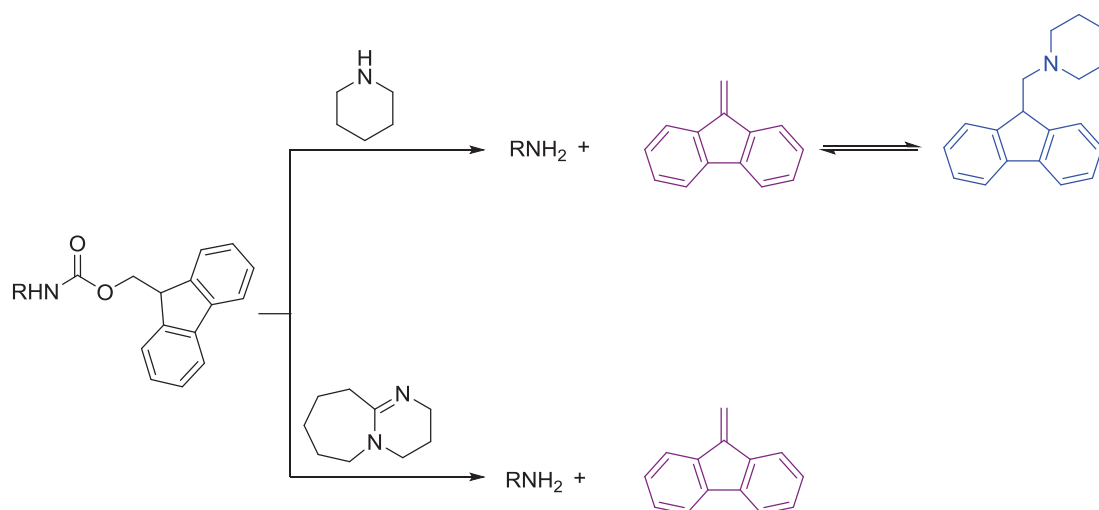
Due to the paramagnetic properties of the NPs, NMR is not available as an analysis method, so alternative strategies had to be used to characterise the loading of the GNPs after each coupling.

Both methods are based on the indirect quantification of sugar coverage (%) from the count of free primary amines on the NPs before and after functionalisation. First, we proceeded with the numbering of the primary amines using the Fmoc method.^[40] This involves adding Fmoc groups onto any free and reactive primary amine available on the NanoLink as received. The removal of the Fmoc group by the DMSO:DMF:DBU (25:24:1) cleavage mixture generated the UV-active dibenzofulvene derivative (Scheme 4.2), which possesses a strong and characteristic absorption at 300 nm and can therefore be quantified by UV spectroscopy.



Scheme 4.2. Reagents and conditions: a) Fmoc chloride, DIPEA, DMF, r.t.; b) DMSO:DMF:DBU (25:24:1), r.t..

This particular cleavage mixture was chosen over the more common cleavage conditions with piperidine to avoid the formation of the piperidine adduct (Scheme 4.3) which, being in equilibrium with the dibenzofulvene but having a different extinction coefficient, might lead to inconsistent or misleading readings, as observed before by previous group members.^[41]



Scheme 4.3. Fmoc group cleavage: after deprotection with piperidine the dibenzofulvene (purple) forms an adduct (blue); when deprotecting with DBU only dibenzofulvene is liberated.

After building a calibration curve by Fmoc cleavage of Fmoc-Gly (using the same conditions), quantifying the free amines available for coupling on the NPs was straightforward. The sugar coverage achieved after each coupling was calculated by subtracting the value obtained for the free amines before and after the decoration of the NPs:

$$\text{coverage \%} = 100 - \left(\frac{Abs_{GNPs}}{Abs_{NanoLink}} \times 100 \right)$$

To compare the results obtained with a different technique we ran a fluorescamine assay.^[42] Fluorescamine itself is non-fluorescent, and only becomes fluorescent after reacting with amines (Fig. 4.4).

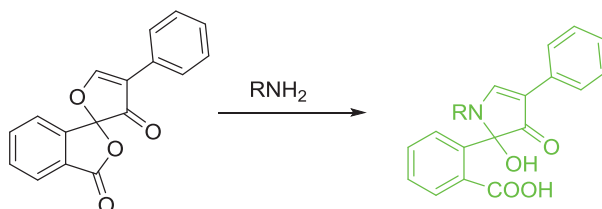


Fig.4.4. Transformation of fluorescamine to the fluorescent adduct by reaction with primary amines.

The fluorescent adduct obtained has its excitation λ at 390 nm and emits at 475 nm. To run the test we simply mixed our NPs with an excess of a fluorescamine solution (calculated from a saturation curve, see SI). Using the same formula as before, the coverage was calculated. As expected, the fluorescent intensity of the NPs treated with fluorescamine decreased as the sugar coverage increased (and fewer free amines were available). This approach is faster than the Fmoc numbering and reproducible within a $\pm 10\%$ correlation level. The main advantage with respect to Fmoc numbering is that Fmoc-Cl could potentially react with the hydroxyl groups of the sugars leading to unreliable measurements, whereas fluorescamine would produce non fluorogenic side products (Fig. 4.5).

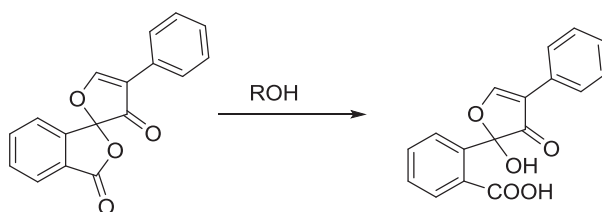


Fig.4.5. Fluorescamine reaction with alcohols: the final adduct is not fluorescent.

4.4 Glyconanoparticle library

4.4.1 Sugar equivalents and coverage

Ideally we aimed to convert all the reactive primary amines available on the NPs with the sugars. Unfortunately, the coverage level of different NPs varied from batch to batch. Aware of the impact this variability might have on the biological testing later on, we tried to establish an average range of coverage common to all the GNPs composing our Glycomap.

We set off to test the reactivity of our library of IME-sugars against the primary amines on NanoLink using the coupling conditions we optimised earlier in this chapter, but varying the equivalents of sugar reagents added. When plotting the loading of the NPs obtained in each reaction against the amount of IME added (Fig. 4.6), it was observed that a plateau was reached for each IME-sugar, suggesting that an even larger excess of reagent would not lead to higher conversion.

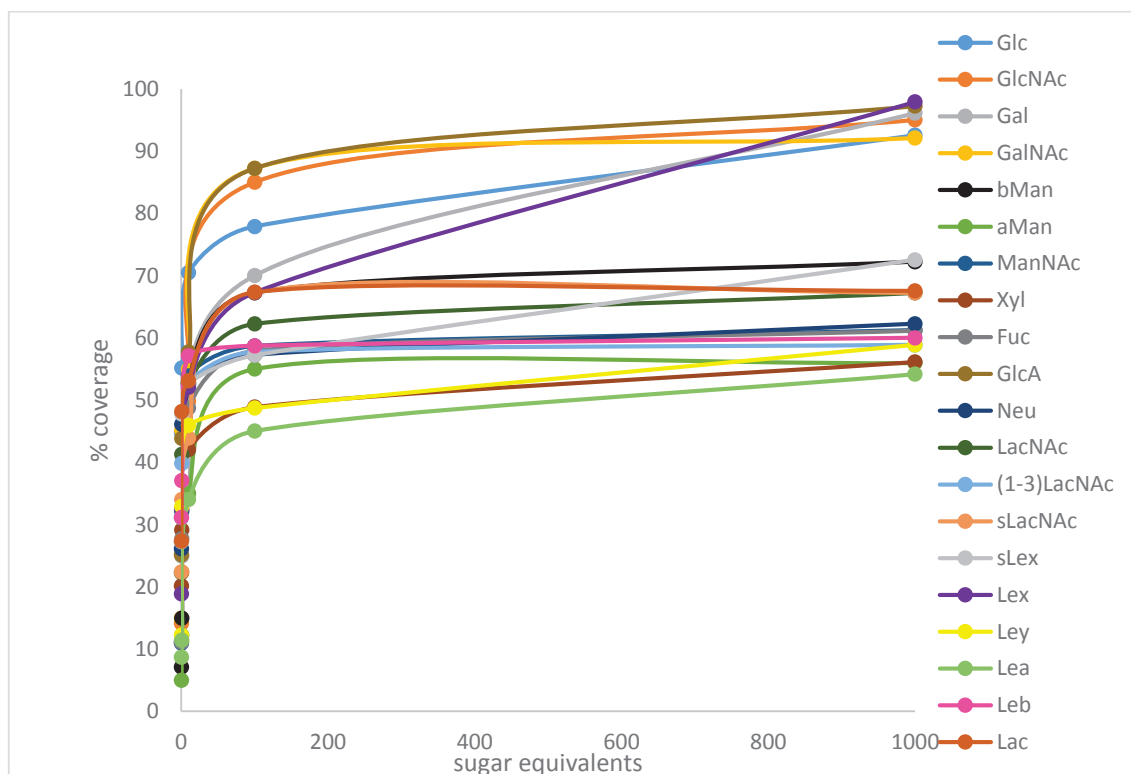


Fig. 4.6 Coverage studies of GNPs. The effect of the progressive increase of sugar-IME equivalents to the NPs on their coverage is reported.

Other considerations can be made from the behaviour of the sugars-IME observed:

- i. Most sugar IMEs are not capable of completely converting the amino functionalities on the NPs, but generally settle around a 60-70% conversion even when using a thousand-fold excess of sugar with respect to the amino groups. This is the case of α and β Man, ManNAc, Xyl, Fuc, Neu, 1-3 and 1,4-LacNAc, sLacNAc, sLex, Le^a, Le^b and Le^y.
- ii. For almost all sugar IMEs the increment in coverage between 0.5 and 10 equivalents is substantial, while it is almost negligible between 100 and 1000 equivalents. Here the only exceptions seem to be Gal and Le^x, which raised the coverage by about 30% (up to saturation) with the last increment. No plateau seems to be visible for these sugars.

- iii. For all sugar IMEs the coverage does not exceed 50% when using between 0.5 and 1 equivalents of reagent.
- iv. A few sugars can reach full coverage. These are Glc, GlcNAc, Gal, GalNAc, GlcA and Le^x.

Overall, no conclusions can be drawn on the reactivity (or inertness) of all these sugar reagents towards the primary amines of the NPs, as their nature is different in terms of size, charge, structure, shape and solubility. A trend could not be easily identified; however it is possible that a mixture of all these factors is contributing to the observed reactivity profile.

Taking in particular the first two considerations into account, we could finally decide how to tailor the library of GNPs feeding our Glycomap, by setting 60-70% of coverage as limiting factor.

4.4.2 Glyconanoparticle library

We synthesised several batches of GNPs until we obtained a coverage of $65\pm 10\%$ for all 20 sugars. In fact, the coverage of the NPs is a crucial factor in the light of biological testing; ^[43,44] therefore, selecting a precise range of coverage is important for consistency during *in vivo* experiments.

The library of GNPs obtained is summarised in the following table:

Sugar	Loading %(Fmoc n)	Loading %(Fluoresc)
β -D-Glc	70	72
β -D-GlcNAc	69	68
β -D-Gal	59	58
β -D-GalNAc	67	66
β -D-Man	61	62
α -D-Man	59	57
α -D-ManNAc	60	60
β -D-Xyl	56	57
α -L-Fuc	65	65
β -D-GlcA	71	70
α -D-Neu	67	66
β -D-Gal-1,4- β -D-Glc	66	68
β -D-Gal-1,4- β -D-GlcNAc	62	62
β -D-Gal-1,3- β -D-GlcNAc	55	55
sialyl-b-D-Gal-1,4- β -D-GlcNAc	66	67
sialyl-Le ^x	72	71
Le ^x	67	66
Le ^y	62	60
Le ^a	52	55
Le ^b	54	56

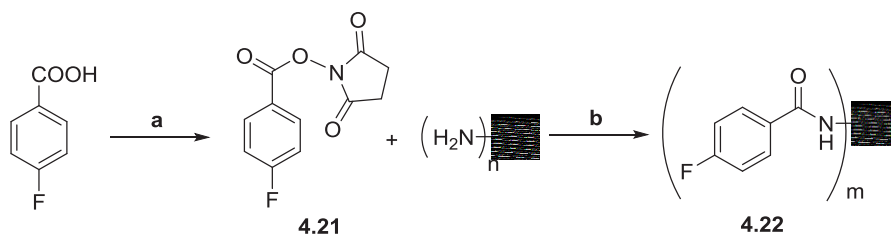
Table 4.1. Library of GNPs. On the left the sugar added to the NPs and on the two columns on the right the coverage obtained.

4.5 Synthesis of [¹⁸F]-GNPs

4.5.1 Synthesis of [¹⁸F]-SFB

Labelling of **4.17** and **4.19** with ¹⁸F required the preparation of [¹⁸F]-N-succinimidyl-4-fluoro-benzoate ([¹⁸F]-SFB), a fluorinating reagent^[45] reacting with primary amines.

Initially we prepared ‘cold’ SFB (Scheme 4.4) in order to test its reactivity towards NanoLink.

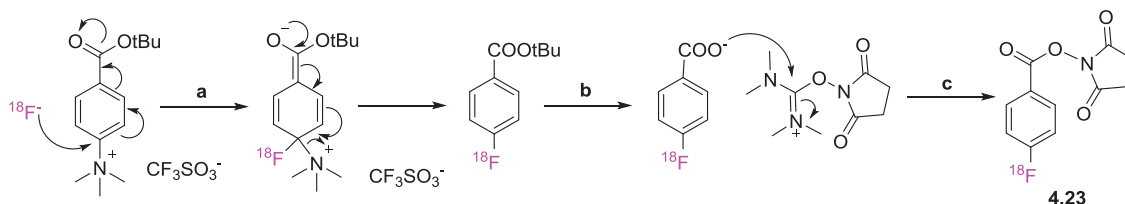


Scheme 4.4. Reagents and conditions: a) NHS, DCC, THF, r.t., 68%; b) DMSO, r.t..

The coverage obtained for the SFB-NPs was 23% (one single batch produced).

Satisfied by this result we proceeded with the synthesis of [^{18}F]-SFB.

The mechanism of this multi-step process involves the substitution of the aromatic ammonium group with ^{18}F and subsequent deprotection allowing the activation of the carboxylic moiety as a NHS ester (Scheme 4.5).



Scheme 4.5. Reagents and conditions: a) K222, K_2CO_3 , ACN, r.t.; b) TPAH, ACN, 120 °C; c) ACN, 90 °C.

All the steps of [^{18}F]-SFB production were carried out inside the hot cell and it was mostly automated by NanoTek (Fig. 4.7).

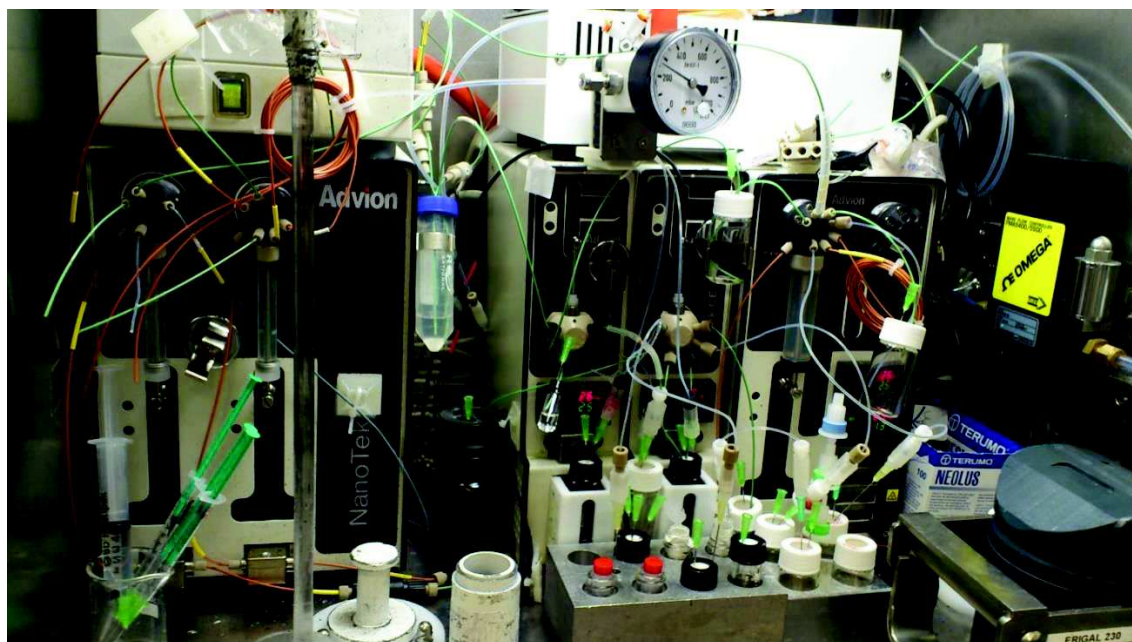


Fig. 4.7. NanoTek after set up. A series of pumps and valves directs the flow of the reagents through subsequent steps in an automated fashion.

A simplified scheme of how NanoTek operates is represented below (Fig. 4.8).

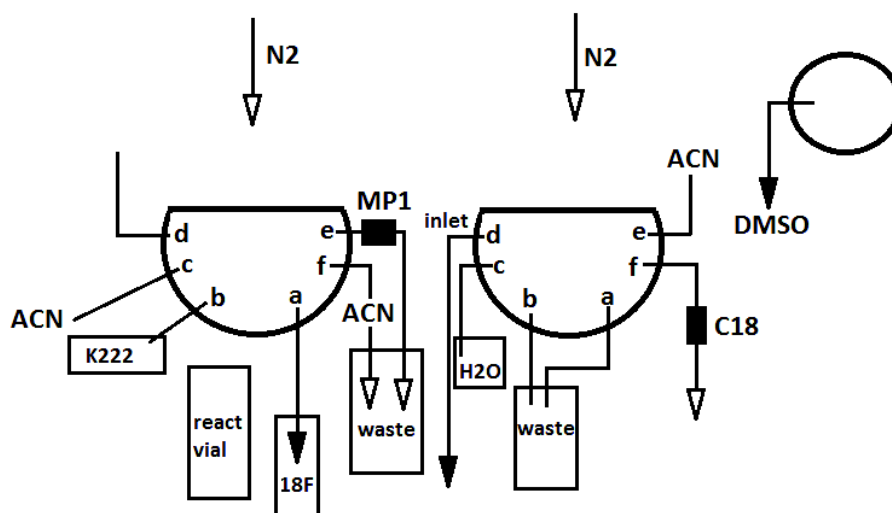
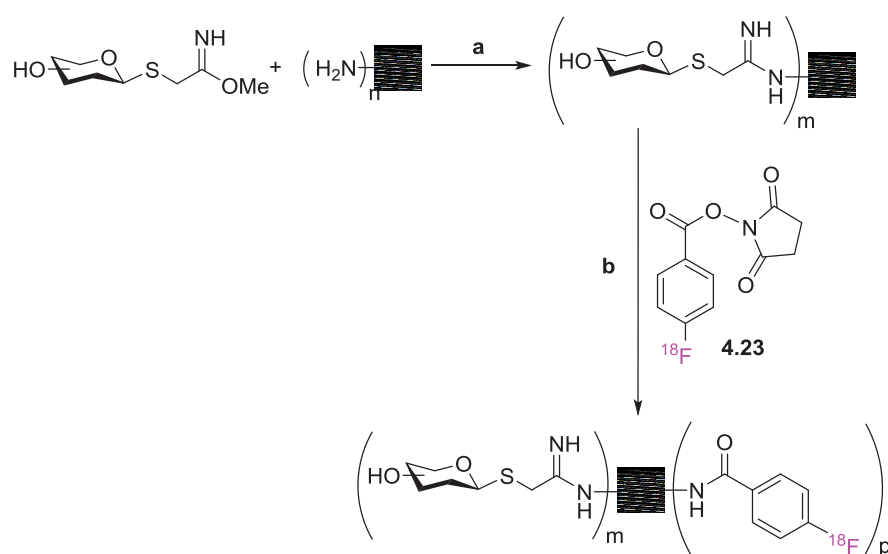


Fig. 4.8. NanoTek workflow. The valves and connections set up in the NanoTek are reported in this schematic diagram.

The detailed procedure, with all the operational steps, is reported in section 4.8.

4.5.2 Synthesis of [^{18}F]-GNPs

Peptide coupling of the activated ester **4.23** has been extensively described in the literature for protein labelling under mild conditions.^[46] Similarly, [^{18}F]-SFB was reacted with the primary amines still available on the surface of Le^x -NPs **4.17** (34%) and Le^a -NPs **4.19** (47%) inside the hot cell (Scheme 4.6).



Scheme 4.6 Reagents and conditions: a) NaHCO_3 0.1 M, pH 9, r.t.; b) DMSO, 30 °C.

The overall process, from receiving ^{18}F , to the synthesis of [^{18}F]-SFB and labelling of the GNPs lasted about 4 h. Even though each step of the hot work was optimised, the activity of the final labelled GNPs was in the order of only few MBq.

This could be due to a number of factors:

- reaction times have been limited to minimise the decay of the radionuclide during the overall process, and this lowered the yield of the chemical reactions;
- the purification steps with C18 and alumina cartridges favour the isolation of a highly pure product (**4.23**) at the cost of losing more reagent and activity;
- the coupling of the radioactive ester to the NPs occurs on an exiguous number of sites available to react on the GNPs and the efficiency of the labelling itself is already limited.

4.6 Conclusions

In this chapter we entered into the details of building the contrasts for the Glycomap. For this purpose we successfully coupled all 20 sugars-IME of the Glycomap in mild aqueous conditions on NanoLink commercial beads, generating a set of GNPs with 65% ($\pm 10\%$) of the reactive functionalities conjugated to homogeneous carbohydrate units. The magneticity of the NPs played an advantage in simplifying the purification steps during the synthesis but forced us to look into alternative characterisation techniques than the ones routinely used in chemistry. We managed to quantify the degree of modification of our GNPs with two different methodologies, by using indirect quantification of the primary amines transformed during the glycosylation step. Fmoc numbering and fluorescamine assay presented very different features in a matter of time required and troubleshooting but both methods gave comparable and reproducible results.

While optimising the conditions for the sugar-IME couplings we explored the possibility of upgrading our platform from MRI to both MRI and PET tracers. After accepting some hard compromises between time constraints and reactivity requirements we managed to label Le^x and Le^a GNPs with ¹⁸F, establishing the first complete protocol for the hot labelling of GNPs.

4.7 References

- [1] N. Blow, *Nat. Methods*. **2009**, 6, 465–469.
- [2] T. F. Massoud, and S. S. Gambhir, *Genes. Dev.* **2003**, 545–580.
- [3] F. Nensa, K. Beiderwellen, P. Heusch, and A. Wetter, *Diagn. Interv. Radiol.* **2014**, 20, 438–447.
- [4] P. McQuade, D. McCarthy, and M. Welch, in *Positron Emission Tomography* (Eds.: D. Bailey, D. Townsend, P. Valk, and M. Maisey), Springer London, **2005**, 237–250.
- [5] S. Kitson, V. Cuccurullo, A. Ciarmiello, D. Salvo, and L. Mansi, *Curr. Radiopharm.* **2009**, 2, 224–253.
- [6] S. Partovi, A. Kohan, C. Rubbert, J. L. Vercher-Conejero, C. Gaeta, R. Yuh, L. Zipp, K. a Herrmann, M. R. Robbin, Z. Lee, R. F. Muzic, P. Faulhaber, and P. R. Ros, *Am. J. Nucl. Med. Mol. Imaging.* **2014**, 4, 202–212.
- [7] Y. Tai and P. Piccini, *J. Neurol. Neurosurg. Psychiatry* **2004**, 75, 669–676.
- [8] G. Antoch, L. S. Freudenberg, T. Beyer, A. Bockisch and J. F. Debatin, *The Journal of Nuclear Medicine* **2004**, 45, 56–66.
- [9] S. Assadian, A. Aliaga, R. F. Del Maestro, A. C. Evans, and B. J. Bedell, *Neuro. Oncol.* **2008**, 10, 292–299.
- [10] P. Seam, M. E. Juweid, and B. D. Cheson, *Blood* **2007**, 110, 3507–16.
- [11] J. Frangioni, *Curr. Opin. Chem. Biol.* **2003**, 7, 626–634.
- [12] P. C. Lauterbur, *Nature* **1973**, 242, 190–191.
- [13] P. Mansfield and P. K. Grannell, *J. Phys. Chem. Solid State Phys.* **1973**, 6, 422–426.
- [14] R. Bronen and G. Sze, *J. Neurosurg.* **1990**, 73, 820–839.

- [15] A. Bjørnerud and L. Johansson, *NMR Biomed.*, 2004, **17**, 465–477.
- [16] M. Pittet, F. Swirski, F. Reynolds, L. Josephson, and R. Weissleder, *Nat. Protoc.* **2006**, 1, 73–79.
- [17] G. J. Strijkers, W. J. M. Mulder, G. A. F. Van Tilborg, and K. Nicolay, *Anticancer Agents Med. Chem.* **2007**, 7, 291–305.
- [18] P. Caravan *Acc. Chem. Res.* **2008**, 42, 851–862.
- [19] S. Iwaki, K. Hokamura, M. Ogawa, Y. Takehara, Y. Muramatsu, T. Yamane, K. Hirabayashi, Y. Morimoto, K. Haggisawa, K. Nakahara, T. Mineno, T. Terai, T. Komatsu, T. Ueno, K. Tamura, Y. Adachi, Y. Hirata, M. Arita, H. Arai, K. Umemura, T. Nagano, and K. Hanaoka, *Org. Biomol. Chem.* **2014**, 12, 8611–8.
- [20] H. Yim, S. Seo, and K. Na, *J. Nanomater.* **2011**, 2011, 1–11.
- [21] K. Chan and W. Wong, *Coord. Chem. Rev.* **2007**, 251, 2428–2451.
- [22] C. Tassa, S. Y. Shaw, and R. Weissleder, *Acc. Chem. Res.* **2011**, 44, 842–852.
- [23] M. L. Mojica Piscioti, E. Lima, M. Vasquez Mansilla, V. E. Tognoli, H. E. Troiani, A. A. Pasa, T. B. Creczynski-Pasa, A. H. Silva, P. Gurman, L. Colombo, G. F. Goya, A. Lamagna, and R. D. Zysler, *J. Biomed. Mater. Res. Part B Appl. Biomater.* **2014**, 102, 860–868.
- [24] J. Li, Y. Hu, J. Yang, W. Sun, H. Cai, P. Wei, Y. Sun, G. Zhang, X. Shi, and M. Shen, *J. Mater. Chem. B* **2015**, 3, 5720–5730.
- [25] R. Weissleder, D. D. Stark, B. L. Engelstad, B. A. Bacon, D. L. White, P. Jacobs, and J. Lewis, *Am. J. Roentgenol.* **1989**, 152, 167–173.
- [26] J. J. Lundquist and E. J. Toone, *Chem. Rev.* **2002**, 102, 555–578.

- [27] i) D. Schwefel, C. Maierhofer, J. G. Beck, S. Seeberger, K. Diederichs, H. M. Möller, W. Welte, and V. Wittmann, *J. Am. Chem. Soc.* **2010**, 132, 8704–8719; ii) M. Mammen, S.-K. Choi, and G. M. Whitesides, *Angew. Chemie Int. Ed.* **1998**, 37, 2754-94.
- [28] S. Lee and X. Chen, *Mol. Imaging* **2009**, 8, 87-100.
- [29] A. Louie, *Chem. Rev.* **2010**, 110, 3146-95.
- [30] i) C. H. Lai, C. Y. Lin, H. T. Wu, H. S. Chan, Y. J. Chuang, C. T. Chen, and C. C. Lin, *Adv. Funct. Mater.* **2010**, 20, 3948–3958; ii) S. Park, G.-H. Kim, S.-H. Park, J. Pai, D. Rathwell, J.-Y. Park, Y.-S. Kang, and I. Shin, *J. Am. Chem. Soc.* **2015**, 137, 5961-8.
- [31] R. Uppal, C. Catana, I. Ay, T. Benner, a G. Sorensen, and P. Caravan, *Radiology* **2011**, 258, 812-20.
- [32] N. R. Sibson, A. M. Blamire, M. Bernades-Silva, S. Laurent, S. Boutry, R. N. Muller, P. Styles, and D. C. Anthony, *Magn. Reson. Med.* **2004**, 51, 248–252; ii) S. I. van Kasteren, S. J. Campbell, S. Serres, D. C. Anthony, N. R. Sibson, and B. G. Davis, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, 106, 18–23.
- [33] B. T. F. Orntoft, E. H. Holmes, P. Johnson, S. Hakomori, and H. Clausen, *Blood* **1991**, 77, 1389-96.
- [34] R. P. McEver, *Glycoconj. J.* **1997**, 14, 585–91.
- [35] G. Bort, D. Anthony, B. G. Davis, *Inflammation Imaging and Therapy*, International patent A61K 49/10 (20060101); A61K 31/702 (20060101); C07H 23/00 (20060101), filed August 15, 2013, and issued August 20, 2015.

- [36] Y. C. Lee, C. P. Stowell, and M. J. Krantz, *Biochemistry* **1976**, 15, 3956–63.
- [37] O. M. T. Pearce, K. D. Fisher, J. Humphries, L. W. Seymour, A. Smith, and B. G. Davis, *Angew. Chem. Int. Ed. Engl.* **2005**, 44, 1057–61.
- [38] M. a Robinson, S. T. Charlton, P. Garnier, X. Wang, S. S. Davis, A. C. Perkins, M. Frier, R. Duncan, T. J. Savage, D. a Wyatt, S. a Watson, and B. G. Davis, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101, 14527–32.
- [39] *The Merck Index, 12th ed., Entry #8726*
- [40] W. Newcomb, T. Deegan, W. Miller, and J. Porco, *Biotechnol. Bioeng.* **1998**, 61, 55–60.
- [41] van Kasteren, S.I. DPhil thesis, University of Oxford: Oxford, **2007**.
- [42] S. Udenfriend, S. Stein, P. Böhlen, W. Dairman, W. Leimgruber, and M. Weigle, *Science* **1972**, 178, 871–872.
- [43] B. Saha, T. H. Evers, and M. W. J. Prins, *Anal. Chem.* **2014**, 86, 8158–66.
- [44] T. J. Cho, J. M. Pettibone, J. M. Gorham, T. M. Nguyen, R. I. MacCuspie, J. Gigault, and V. a. Hackley, *Langmuir* **2015**, 31, 7673–7683.
- [45] U. Ackermann, S. D. Yeoh, J. I. Sachinidis, S. S. Poniger, A. M. Scott, and H. J. Tochon-Danguy, *J. Label. Compd. Radiopharm.* **2011**, 54, 671–673.
- [46] G. Tang, W. Zeng, M. Yu, and G. Kabalka, *J. Label. Compd. Radiopharm.* **2008**, 51, 68–71.

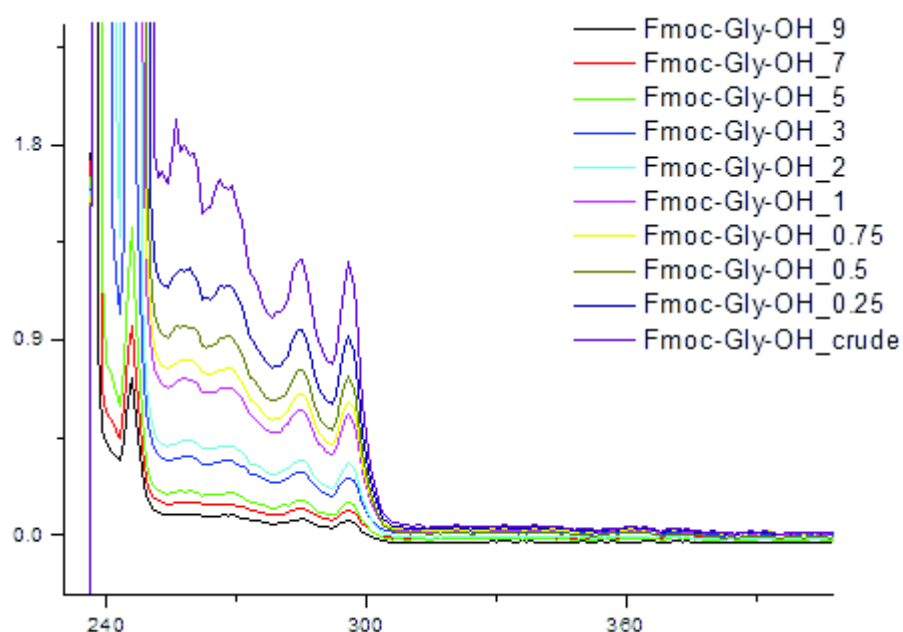
4.8 Supporting information

Fmoc numbering protocol

Fmoc-Gly-OH (116.8 mg, 11.6 μmol) was treated with DMSO:DMF:DBU (25:24:1 solution, 65 mL) at r.t. to cleave the Fmoc group. After 2 h, TLC (EtOAc:MeOH, 8:2) indicated the formation of the product ($R_f = 0.0$) with the complete consumption of Fmoc-Gly-OH ($R_f = 0.2$). From the crude, 10 samples were prepared to build a calibration curve (Fig. S1):

- 1) 100 μL crude
- 2) 100 μL crude + 25 μL of DMSO:DMF:DBU (25:24:1 solution)
- 3) 100 μL crude + 50 μL of DMSO:DMF:DBU (25:24:1 solution)
- 4) 100 μL crude + 75 μL of DMSO:DMF:DBU (25:24:1 solution)
- 5) 100 μL crude + 100 μL of DMSO:DMF:DBU (25:24:1 solution)
- 6) 100 μL crude + 200 μL of DMSO:DMF:DBU (25:24:1 solution)
- 7) 100 μL crude + 300 μL of DMSO:DMF:DBU (25:24:1 solution)
- 8) 100 μL crude + 500 μL of DMSO:DMF:DBU (25:24:1 solution)
- 9) 100 μL crude + 700 μL of DMSO:DMF:DBU (25:24:1 solution)
- 10) 100 μL crude + 900 μL of DMSO:DMF:DBU (25:24:1 solution)

The samples were diluted with MeOH (20 μL sample + 1000 μL MeOH) and the absorbance measured by UV.



conc (M)	Abs
1.184E-04	1.260553
9.472E-05	0.916939
7.893E-05	0.732927
6.766E-05	0.617422
5.920E-05	0.556130
3.947E-05	0.334115
2.960E-05	0.266010
1.973E-05	0.151513
1.480E-05	0.104092
1.184E-05	0.067799

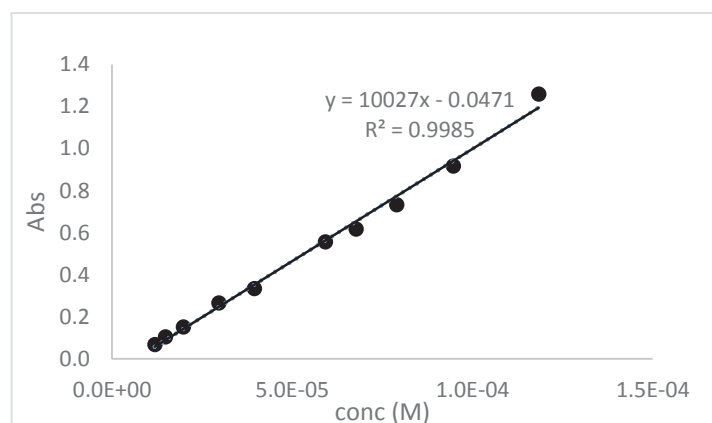


Figure S1. Fmoc numbering calibration curve built from progressive dilution of Fmoc-Gly cleavage mixture.

The ϵ obtained at 295 nm was $10,027 \text{ cm}^{-1}\text{M}^{-1}$.

To NanoLink (0.1 mg, Amino Magnetic Beads $0.8 \mu\text{m}$, 41.26 nmol/mg of NH_2 , for a total of $4.6 \mu\text{mol}$ of NH_2 , 1 eq.) in DMF ($100 \mu\text{L}$) at 0°C were added Fmoc-Cl (2.9 mg, $11 \mu\text{mol}$, 2.4 eq.) and DIPEA ($3.4 \mu\text{L}$, $19 \mu\text{mol}$, 4.2 eq.). The reaction mixture was stirred at r.t. for 16 h. The particles were then separated using a magnet and washed repeatedly with DMF. The cleavage mixture

(DMSO:DMF:DBU, 25:24:1, 500 μ L) was used to re-suspend the NPs and the mixture was stirred at r.t. for 2 h. NPs were again separated with a magnet, 20 μ L of the solution withdrawn and added to 1 ml of MeOH before measuring at 295 nm. Blank was measured using 20 μ L of cleavage mixture in 1 mL of MeOH.

$$\frac{mmol}{g} = \left(\left(\left(\frac{A}{\epsilon b} \right) \times 100 \right) \times 0.5 \right) \times 1000$$

GNPs were measured using the same protocol, adopting as reference the coverage obtained for NanoLink.

Fluorescamine assay protocol

From a stock solution of fluorescamine (2.5 mg in 500 μ L of DMSO) was prepared a series of fluorescamine solutions in progressive dilutions:

- A. 0.010 mg/mL (250 μ L solution B + 250 μ L DMSO)
- B. 0.020 mg/mL (250 μ L solution C + 250 μ L DMSO)
- C. 0.039 mg/mL (250 μ L solution D + 250 μ L DMSO)
- D. 0.078 mg/mL (250 μ L solution E + 250 μ L DMSO)
- E. 0.156 mg/mL (250 μ L solution F + 250 μ L DMSO)
- F. 0.312 mg/mL (250 μ L solution G + 250 μ L DMSO)
- G. 0.625 mg/mL (250 μ L solution H + 250 μ L DMSO)
- H. 1.25 mg/mL (250 μ L solution I + 250 μ L DMSO)
- I. 2.5 mg/mL (250 μ L stock solution + 250 μ L DMSO)
- J. 5 mg/mL (250 μ L of stock solution)

From a stock solution of NanoLink (0.5 mg in 3 mL of PBS buffer, 50 mM, pH 8) were prepared 11 samples to build a saturation curve (Fig.S2):

1. 150 μ L NanoLink + 50 μ L DMSO
2. 150 μ L NanoLink + 50 μ L fluorescamine solution A
3. 150 μ L NanoLink + 50 μ L fluorescamine solution B
4. 150 μ L NanoLink + 50 μ L fluorescamine solution C
5. 150 μ L NanoLink + 50 μ L fluorescamine solution D
6. 150 μ L NanoLink + 50 μ L fluorescamine solution E
7. 150 μ L NanoLink + 50 μ L fluorescamine solution F
8. 150 μ L NanoLink + 50 μ L fluorescamine solution G
9. 150 μ L NanoLink + 50 μ L fluorescamine solution H
10. 150 μ L NanoLink + 50 μ L fluorescamine solution I
11. 150 μ L NanoLink + 50 μ L fluorescamine solution J

After shaking the samples at r.t. for 15 min the absorbance was read using a plate reader (15 μ L sample per dwell, excitation λ = 350 nm, emission λ = 460 nm).

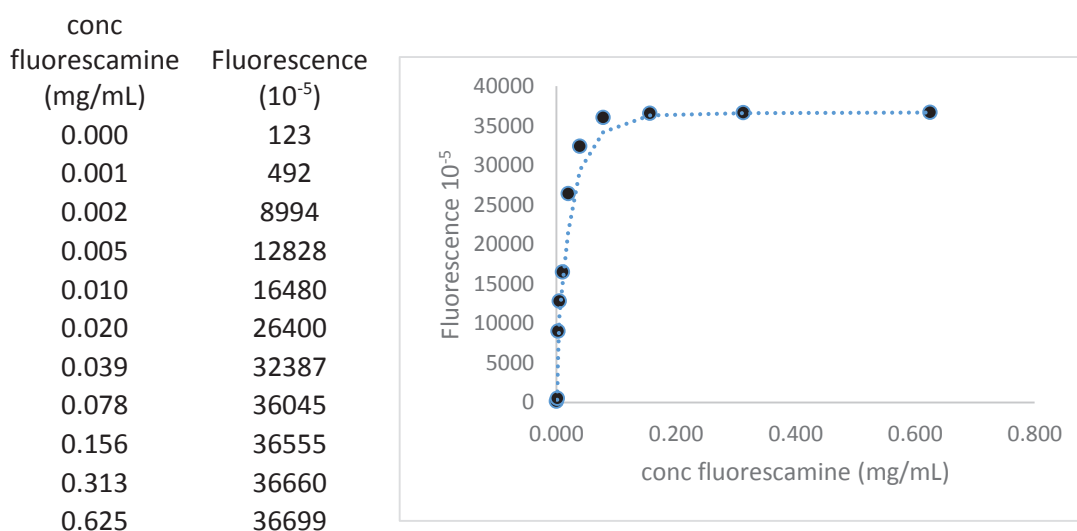


Figure S2. Saturation curve built from the fluorescent response of NanoLink to addition of fluorescamine solution at increasing concentration.

To measure the samples a fluorescamine stock solution (50 μ L, 2.5 mg in 500 μ L of DMSO) was added to the NPs (150 μ L, 0.05 mg in 300 μ L of phosphate buffer, 10 mM, pH 7.4). After 15 min shaking at r.t. the sample was measured. Background was measured using stock solution of NanoLink (150 μ L, 0.05 mg in 300 μ L of phosphate buffer, 10 mM, pH 7.4) and DMSO (50 μ L) and blank was measured using a mixture of phosphate buffer (150 μ L, 10 mM, pH 7.4) and DMSO (50 μ L). GNPs coverage was obtained using NanoLink as reference.

Sugars equivalents and NPs coverage

To NanoLink-amino magnetic beads (5 x 100 μ L, 10 mg/ml, 5 x 1 mg, 5 x 0.041 μ mol-NH₂, 1 eq.), previously washed with milliQ water and recovered with NaHCO₃ (0.1 M, 500 μ L), were added sugars-IME (0.021, 0.041, 0.41, 4.1, 41 μ mol, 0.5, 1, 10, 100, 1000 eq.) and the reaction was stirred for 16 h at r.t.. The particles were separated by magnet and washed with NaHCO₃ (0.1 M, 3 x 500 μ L). The functionalisation was quantified by fluorescamine assay and Fmoc numbering.

Fluorescamine assay (Fig. S3):

Row	Well	Sample	Crude data	Blank corrected data (350, 460)	% coverage
P	1	Blank B	156723		
P	2	Standard S1	189342	32619	0.00
P	3	0.5 eq-Glc	174763	18040	44.69
P	4	1 eq-Glc	171301	14578	55.31
P	5	10 eq-Glc	166208	9485	70.92
P	6	100 eq-Glc	163799	7076	78.31

P	7	1000 eq-Glc	159132	2409	92.61
A	1	0.5 eq-GlcNAc	184575	27852	14.61
A	2	1 eq-GlcNAc	182065	25342	22.31
A	3	10 eq-GlcNAc	173132	16409	49.69
A	4	100 eq-GlcNAc	161605	4882	85.03
A	5	1000 eq-GlcNAc	158333	1610	95.06
B	1	0.5 eq-Gal	185723	29000	11.09
B	2	1 eq-Gal	181207	24484	24.94
B	3	10 eq-Gal	177895	21172	35.09
B	4	100 eq-Gal	166498	9775	70.03
B	5	1000 eq-Gal	158037	1314	95.97
C	1	0.5 eq-GalNAc	184450	27727	15.00
C	2	1 eq-GalNAc	174563	17840	45.31
C	3	10 eq-GalNAc	171411	14688	54.97
C	4	100 eq-GalNAc	160973	4250	86.97
C	5	1000 eq-GalNAc	159432	2709	91.70
D	1	0.5 eq-bMan	187158	30435	6.70
D	2	1 eq-bMan	184459	27736	14.97
D	3	10 eq-bMan	172153	15430	52.70
D	4	100 eq-bMan	167477	10754	67.03
D	5	1000 eq-bMan	165866	9143	71.97
E	1	0.5 eq-aMan	187701	30978	5.03
E	2	1 eq-aMan	185438	28715	11.97
E	3	10 eq-aMan	177915	21192	35.03
E	4	100 eq-aMan	171401	14678	55.00
E	5	1000 eq-aMan	171175	14452	55.69
F	1	0.5 eq-ManNAc	185853	29130	10.70
F	2	1 eq-ManNAc	178983	22260	31.76
F	3	10 eq-ManNAc	171737	15014	53.97
F	4	100 eq-ManNAc	170196	13473	58.70
F	5	1000 eq-ManNAc	169344	12621	61.31
G	1	0.5 eq-Xyl	182718	25995	20.31
G	2	1 eq-Xyl	179682	22959	29.61
G	3	10 eq-Xyl	175752	19029	41.66
G	4	100 eq-Xyl	173348	16625	49.03
G	5	1000 eq-Xyl	171064	14341	56.03
H	1	0.5 eq-Fuc	185752	29029	11.01
H	2	1 eq-Fuc	180210	23487	28.00
H	3	10 eq-Fuc	173458	16735	48.70
H	4	100 eq-Fuc	170322	13599	58.31
H	5	1000 eq-Fuc	169434	12711	61.03
I	1	0.5 eq-GlcA	181087	24364	25.31
I	2	1 eq-GlcA	174984	18261	44.02
I	3	10 eq-GlcA	170424	13701	58.00
I	4	100 eq-GlcA	160969	4246	86.98
I	5	1000 eq-GlcA	157710	987	96.97

J	1	0.5 eq-Neu	180860	24137	26.00
J	2	1 eq-Neu	174330	17607	46.02
J	3	10 eq-Neu	172047	15324	53.02
J	4	100 eq-Neu	170744	14021	57.02
J	5	1000 eq-Neu	169111	12388	62.02
K	1	0.5 eq-(1,4)-LacNAc	182155	25432	22.03
K	2	1 eq-(1,4)-LacNAc	175868	19145	41.31
K	3	10 eq-(1,4)-LacNAc	172153	15430	52.70
K	4	100 eq-(1,4)-LacNAc	169108	12385	62.03
K	5	1000 eq-(1,4)-LacNAc	167497	10774	66.97
L	1	0.5 eq-(1,3)-LacNAc	179230	22507	31.00
L	2	1 eq-(1,3)-LacNAc	176394	19671	39.69
L	3	10 eq-(1,3)-LacNAc	172059	15336	52.98
L	4	100 eq-(1,3)-LacNAc	170429	13706	57.98
L	5	1000 eq-(1,3)-LacNAc	170100	13377	58.99
M	1	0.5 eq-sLacNAc	182065	25342	22.31
M	2	1 eq-sLacNAc	178151	21428	34.31
M	3	10 eq-sLacNAc	174979	18256	44.03
M	4	100 eq-sLacNAc	167467	10744	67.06
M	5	1000 eq-sLacNAc	167497	10774	66.97
N	1	0.5 eq-sLex	180534	23811	27.00
N	2	1 eq-sLex	173664	16941	48.06
N	3	10 eq-sLex	172073	15350	52.94
N	4	100 eq-sLex	170779	14056	56.91
N	5	1000 eq-sLex	165656	8933	72.61
O	1	0.5 eq-Lex	183144	26421	19.00
O	2	1 eq-Lex	178703	21980	32.62
O	3	10 eq-Lex	172480	15757	51.69
O	4	100 eq-Lex	167461	10738	67.08
O	5	1000 eq-Lex	157275	552	98.31
P	1	0.5 eq-Ley	185427	28704	12.00
P	2	1 eq-Ley	178477	21754	33.31
P	3	10 eq-Ley	174437	17714	45.69
P	4	100 eq-Ley	173458	16735	48.70
P	5	1000 eq-Ley	170096	13373	59.00
A2	1	0.5 eq-Lea	186506	29783	8.69
A2	2	1 eq-Lea	185553	28830	11.62
A2	3	10 eq-Lea	178151	21428	34.31
A2	4	100 eq-Lea	174633	17910	45.09
A2	5	1000 eq-Lea	171527	14804	54.62
B2	1	0.5 eq-Leb	179210	22487	31.06
B2	2	1 eq-Leb	177372	20649	36.70
B2	3	10 eq-Leb	170849	14126	56.69
B2	4	100 eq-Leb	170196	13473	58.70
B2	5	1000 eq-Leb	169770	13047	60.00
B3	1	0.5 eq-Lac	188790	16161	27.26

B3	2	1 eq-Lac	183992	11363	48.85
B3	3	10 eq-Lac	182900	10271	53.77
B3	4	100 eq-Lac	179925	7296	67.16
B3	5	1000 eq-Lac	179825	7196	67.61

Figure S3. All reported values of fluorescence for GNPs fluorescamine assay: coverage is obtained from the neat fluorescence value (after blank subtraction).

Fmoc numbering (Fig. S4):

Sample	Abs	conc (M)	Mmol	% coverage
NanoLink	0.80	7.97846E-05	0.040	0
0.5 eq-Glc	0.44	4.38815E-05	0.022	45
1 eq-Glc	0.36	3.59031E-05	0.018	55
10 eq-Glc	0.24	2.39354E-05	0.012	70
100 eq-Glc	0.18	1.79515E-05	0.009	78
1000 eq-Glc	0.06	5.98384E-06	0.003	93
0.5 eq-GlcNAc	0.69	6.88142E-05	0.034	14
1 eq-GlcNAc	0.62	6.18331E-05	0.031	23
10 eq-GlcNAc	0.40	3.98923E-05	0.020	50
100 eq-GlcNAc	0.12	1.19677E-05	0.006	85
1000 eq-GlcNAc	0.04	3.98923E-06	0.002	95
0.5 eq-Gal	0.71	7.08088E-05	0.035	11
1 eq-Gal	0.60	5.98384E-05	0.030	25
10 eq-Gal	0.52	5.186E-05	0.026	35
100 eq-Gal	0.24	2.39354E-05	0.012	70
1000 eq-Gal	0.03	2.99192E-06	0.001	96
0.5 eq-GalNAc	0.68	6.78169E-05	0.034	15
1 eq-GalNAc	0.44	4.38815E-05	0.022	45
10 eq-GalNAc	0.36	3.59031E-05	0.018	55
100 eq-GalNAc	0.10	9.97307E-06	0.005	88
1000 eq-GalNAc	0.06	5.98384E-06	0.003	93
0.5 eq-bMan	0.74	7.38007E-05	0.037	8
1 eq-bMan	0.68	6.78169E-05	0.034	15
10 eq-bMan	0.38	3.78977E-05	0.019	53
100 eq-bMan	0.26	2.593E-05	0.013	68
1000 eq-bMan	0.22	2.19408E-05	0.011	73
0.5 eq-aMan	0.76	7.57954E-05	0.038	5
1 eq-aMan	0.70	6.98115E-05	0.035	13
10 eq-aMan	0.52	5.186E-05	0.026	35
100 eq-aMan	0.36	3.59031E-05	0.018	55
1000 eq-aMan	0.35	3.49058E-05	0.017	56

0.5 eq-ManNAc	0.71	7.08088E-05	0.035	11
1 eq-ManNAc	0.54	5.38546E-05	0.027	33
10 eq-ManNAc	0.37	3.69004E-05	0.018	54
100 eq-ManNAc	0.33	3.29111E-05	0.016	59
1000 eq-ManNAc	0.31	3.09165E-05	0.015	61
0.5 eq-Xyl	0.64	6.38277E-05	0.032	20
1 eq-Xyl	0.57	5.68465E-05	0.028	29
10 eq-Xyl	0.46	4.58761E-05	0.023	43
100 eq-Xyl	0.41	4.08896E-05	0.020	49
1000 eq-Xyl	0.35	3.49058E-05	0.017	56
0.5 eq-Fuc	0.71	7.08088E-05	0.035	11
1 eq-Fuc	0.58	5.78438E-05	0.029	28
10 eq-Fuc	0.41	4.08896E-05	0.020	49
100 eq-Fuc	0.34	3.39084E-05	0.017	58
1000 eq-Fuc	0.31	3.09165E-05	0.015	61
0.5 eq-GlcA	0.60	5.98384E-05	0.030	25
1 eq-GlcA	0.45	4.48788E-05	0.022	44
10 eq-GlcA	0.34	3.39084E-05	0.017	58
100 eq-GlcA	0.10	9.97307E-06	0.005	88
1000 eq-GlcA	0.02	1.99461E-06	0.001	98
0.5 eq-Neu	0.59	5.88411E-05	0.029	26
1 eq-Neu	0.43	4.28842E-05	0.021	46
10 eq-Neu	0.38	3.78977E-05	0.019	53
100 eq-Neu	0.34	3.39084E-05	0.017	58
1000 eq-Neu	0.30	2.99192E-05	0.015	63
0.5 eq-(1,4)-LacNAc	0.62	6.18331E-05	0.031	23
1 eq-(1,4)-LacNAc	0.47	4.68734E-05	0.023	41
10 eq-(1,4)-LacNAc	0.38	3.78977E-05	0.019	53
100 eq-(1,4)-LacNAc	0.30	2.99192E-05	0.015	63
1000 eq-(1,4)-LacNAc	0.26	2.593E-05	0.013	68
0.5 eq-(1,3)-LacNAc	0.55	5.48519E-05	0.027	31
1 eq-(1,3)-LacNAc	0.48	4.78707E-05	0.024	40
10 eq-(1,3)-LacNAc	0.38	3.78977E-05	0.019	53
100 eq-(1,3)-LacNAc	0.34	3.39084E-05	0.017	58
1000 eq-(1,3)-LacNAc	0.33	3.29111E-05	0.016	59
0.5 eq-sLacNAc	0.62	6.18331E-05	0.031	23
1 eq-sLacNAc	0.53	5.28573E-05	0.026	34
10 eq-sLacNAc	0.45	4.48788E-05	0.022	44
100 eq-sLacNAc	0.26	2.593E-05	0.013	68
1000 eq-sLacNAc	0.26	2.593E-05	0.013	68
0.5 eq-sLex	0.58	5.78438E-05	0.029	28
1 eq-sLex	0.42	4.18869E-05	0.021	48
10 eq-sLex	0.38	3.78977E-05	0.019	53
100 eq-sLex	0.34	3.39084E-05	0.017	58
1000 eq-sLex	0.22	2.19408E-05	0.011	73
0.5 eq-Lex	0.65	6.4825E-05	0.032	19

1 eq-Lex	0.54	5.38546E-05	0.027	33
10 eq-Lex	0.38	3.78977E-05	0.019	53
100 eq-Lex	0.26	2.593E-05	0.013	68
1000 eq-Lex	0.02	1.99461E-06	0.001	98
0.5 eq-Ley	0.70	6.98115E-05	0.035	13
1 eq-Ley	0.54	5.38546E-05	0.027	33
10 eq-Ley	0.43	4.28842E-05	0.021	46
100 eq-Ley	0.41	4.08896E-05	0.020	49
1000 eq-Ley	0.33	3.29111E-05	0.016	59
0.5 eq-Lea	0.73	7.28034E-05	0.036	9
1 eq-Lea	0.71	7.08088E-05	0.035	11
10 eq-Lea	0.53	5.28573E-05	0.026	34
100 eq-Lea	0.44	4.38815E-05	0.022	45
1000 eq-Lea	0.37	3.69004E-05	0.018	54
0.5 eq-Leb	0.55	5.48519E-05	0.027	31
1 eq-Leb	0.50	4.98654E-05	0.025	38
10 eq-Leb	0.34	3.39084E-05	0.017	58
100 eq-Leb	0.33	3.29111E-05	0.016	59
1000 eq-Leb	0.32	3.19138E-05	0.016	60
0.5 eq-Lac	0.58	5.78438E-05	0.029	28
1 eq-Lac	0.42	4.18869E-05	0.021	48
10 eq-Lac	0.38	3.78977E-05	0.019	53
100 eq-Lac	0.26	2.593E-05	0.013	68
1000 eq-Lac	0.26	2.593E-05	0.013	68

Figure S4. All reported values of UV absorbance for GNPs Fmoc-numbering assay: coverage is obtained from the absorbance value.

GNPs synthesis

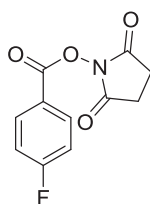
To NanoLink-amino magnetic beads (300 μ L, 10 mg/ml, 3 mg, 0.124 μ mol-NH₂, 1.0 eq.), previously washed with milliQ water and recovered with NaHCO₃ (0.1 M, 500 μ L), was added the IME sugar mixture and the reaction was stirred for 16 h at r.t.. The particles were separated by magnet and washed with NaHCO₃ (0.1 M, 3 x 500 μ L). The functionalisation was quantified by fluorescamine assay and Fmoc numbering (Fig. S5).

Entry	product	Quantity	Coverage (Fluor.)	Coverage (Fmoc)
4.1		79 μmol (613 eq)	71%	70%
4.2		42 μmol (326 eq)	58%	59%
4.3		69 μmol (554 eq)	57%	59%
4.4		6.5 μmol (52 eq)	62%	61%
4.5		3.4 μmol (27 eq)	68%	69%
4.6		91 μmol (734 eq)	66%	67%
4.7		38 μmol (295 eq)	60%	60%
4.8		14 μmol (113 eq)	57%	56%
4.9		9 μmol (72 eq)	65%	65%
4.10		11.5 μmol (93 eq)	70%	70%
4.11		3.9 μmol (32 eq)	66%	67%
4.12		46 μmol (371 eq)	68%	66%
4.13		9.2 μmol (74 eq)	62%	62%
4.14		3.7 μmol (30 eq)	55%	55%
4.15		7.1 μmol (57 eq)	67%	66%
4.16		5.6 μmol (45 eq)	72%	71%
4.17		4.4 μmol (35 eq)	66%	67%

4.18		7.9 μ mol (64 eq)	60%	62%
4.19		4.6 μ mol (37 eq)	55%	52%
4.20		8.0 μ mol (64 eq)	56%	54%

Figure S5. The GNPs library: compound number, final structure, sugar-IME equivalents and coverage obtained.

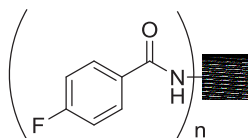
N-succinimidyl-4-F-benzoate **4.21**



To a solution of 4-F-benzoic acid (1.3 g, 9.3 mmol, 1 eq.) in THF (15 mL), were added NHS (1.287 g, 11.2 mmol, 1.2 eq.) and DCC (1.827 g, 8.85 mmol, 0.95 eq). The reaction was stirred at r.t. for 16 h, filtered on celite, evaporated and purified by flash column chromatography on silica (DCM 100%), to afford the title compound **4.21** as a white solid (1.507 g, 6.35 mmol, 68%). **¹H NMR** (CDCl₃, 200 MHz) δ 8.18 (dd, J = 9.0, 5.3 Hz, 2 H, 2 x CHCC), 7.22 (dd, J = 15.1, 6.6 Hz, 2 H, 2 x CHCF), 2.92 (s, 4 H, NHS); **¹³C NMR** (CDCl₃, 126 MHz) δ 169.89 (CF), 169.62 (2 x CON), 166.80 (COO), 161.37 (CCO), 133.95, 133.76 (2 x CHCCO),

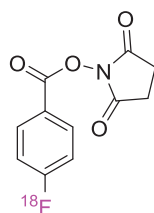
116.99, 116.55 (2 x CHCF), 26.11 (2 x CH₂); **MS** (ESI⁻) *m/z* = 236.17 [M-H]; **IR** (ATR) 1723.72 (CO).

N-succinimidyl-4-F-benzoyl-NanoLink 4.22



To NanoLink-amino magnetic beads (100 μ L, 10 mg/ml, 1 mg, 0.041 μ mol-NH₂, 1 eq.), previously washed with milliQ water and recovered with DMSO (500 μ L), was added **4.21** (6.5 mg, 27.6 μ mol, 674 eq.) and the reaction was stirred for 1 h at r.t.. The particles were separated by magnet and washed with DMSO (500 μ L) and milliQ H₂O (3 x 500 μ L). The functionalisation was quantified by fluorescamine assay (obtained = 25% coverage) and Fmoc numbering (obtained = 22% coverage).

[¹⁸F]-N-succinimidyl-4-F-benzoate 4.23



The fluoride drying was carried out with a NanoTek (Advion BioSystems) device. Cyclotron-produced non-carrier-added aqueous ¹⁸F⁻ (3-5 GBq in 5-6 mL), produced by bombardment of [¹⁸O]-H₂O with protons, was first trapped in an MP1

anion-exchange cartridge and subsequently released into the reactor with Kryptofix® 222 (12 mg/mL) and K₂CO₃ (2.2 mg/mL) in ACN:H₂O (4:1, 5 × 100 µL). The solvent was then evaporated under a stream of nitrogen (0.4 L/min) at 105 °C. The fluoride was further dried azeotropically with ACN (6 × 50 µL). The total drying process took 15-20 min.

The synthesis of **4.23** was performed in assembled NanoTek Liquid Flow reactor. 4-(tert-butoxycarbonylmethyl)phenyl trimethylammonium trifluoromethanesulphonate (4.2 mg, 11 µmol, 1 eq.) in ACN (0.5 mL) was introduced into the reactor. The mixture was heated to 120 °C before the addition of TPAH (1.0 M in H₂O, 20 µL, 20 µmol, 1.8 eq.) in ACN (0.5 mL). The mixture was stirred for 3 min at 120 °C before the solvents were evaporated under a stream of nitrogen. The residue was cooled to 90 °C under N₂ flow and azeotropically dried by addition of ACN (10 × 100 µL). N,N,N',N'-tetramethyl-O-(N-succinimidyl)uronium tetrafluoroborate (12 mg, 40 µmol, 3.6 eq.) in ACN (dry, 1 mL) was added to the reactor and the mixture was stirred at 90 °C for 10 min. The mixture was then cooled to 25 °C and diluted with AcOH (5%, 15 mL). The reaction crude was then transferred through a SEP-PAK plus C18 cartridge. The cartridge was washed with H₂O (5 × 1 mL) and ACN (15% in H₂O), dried by passing air through. The SEP-PAK plus C18 cartridge was then connected to a QMA and 2 successive SEP-PAK light alumina cartridges. The product was finally eluted with Et₂O (1.3 mL), dried under N₂ flow to afford the title compound **4.23** in 34% yield (decay-corrected) with a radiochemical purity exceeding 95% by HPLC analysis.

HPLC method

The HPLC analyses were carried out on UltiMate 3000 System (Dionex) with a reverse phase column C18 Nova-Pak® WAT086344 (250 x 0.18 mm, stationary phase: 4 µm) using an UltiMate 3000 detector (Dionex) at the wavelength of 254 nm and a radiation detector (LabLogic). The elution was effected by the linear gradient of (A = 0.1% TFA/H₂O, B = 0.1% TFA/ACN).

0-1 min 5% B isocratic

1-11 min (5% B to 95% B) linear increase

11-13 min (95% B) isocratic

13-14 min (95% B to 5% B) linear decrease

14-16 min (5% B) isocratic

Flow rate was 1 mL/min. The temperature of the column was 25 °C. The volume of injection was 20 µL. Data were treated on Chromeleon v6.8 (Fig. S6).

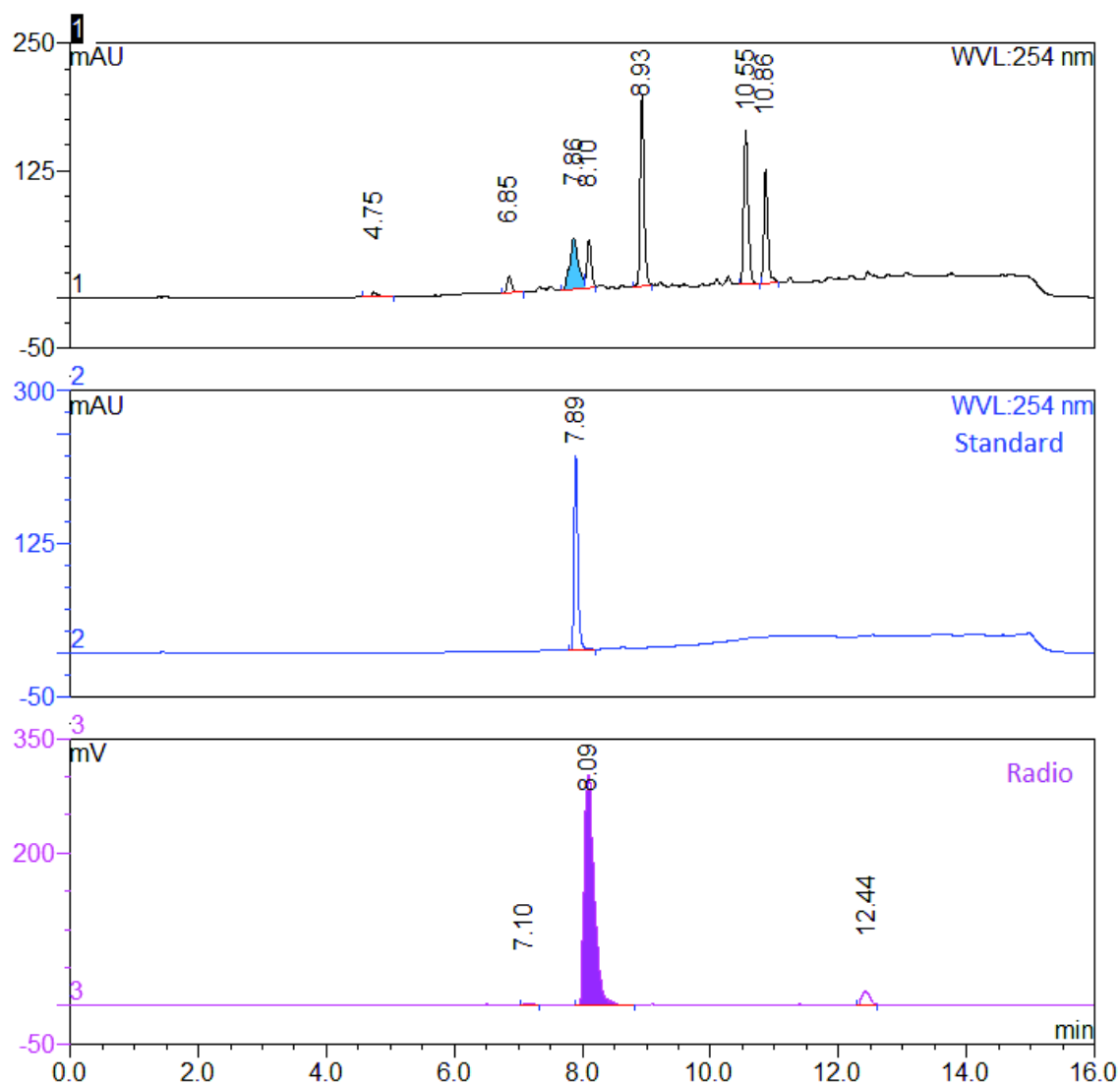


Figure S6. From the top: HPLC trace of the crude reaction mixture, purified compound and radioactive trace.

Operational steps

i. Initial steps

The aspiration, HPLC and laptop are turned on.

NanoTek is switched on in HC1 and the program started

- initialize

- re-initialize

Starting sequence - right click

- load list
- batch+CE2 purification (last sequence)
- load

Turn on DH-1 and CE-1	concentration	0.3 - 2 - 0.3 (start stirring)
	mass flow	2 mL/min

ii. Fluoride drying

- open pot
- take out vial
- measure activity
- put inlet (+vent)
- start program 1 (+radioactivity sensor on MP1) - (tilt ^{18}F vial)
- move MP1 exit to reaction vial (+radioactivity sensor)
- push up vial with K222+K₂CO₃ solution
- remove MP1
- add line f (ACN) to reaction vial

iii. [^{18}F]-SFB synthesis and purification

- take all off from reaction vial
- change T to 90°C (CE1)
- measure activity
- start stirring of reaction vial and add precursor
- leave to react for 10 min

- change T to 120 °C
- add TPAH
- leave to react for 3 min
- take out reaction vial and vent with needle by tapping quickly (leave vent on)
- put back reaction vial to stir, add N₂
- move f in reaction vial

- (computer) conc1
 c-ACN
 absolute
 1000
 9999 - OK

conc1	} (x 5 times)
f-pick up	
dispense	
100	
1000 - OK	

- change T to 90 °C
- (computer) conc1
 f-pick up
 dispense
 100
 1000 - OK
- take off everything from reaction vial

- add coupling reagent
- leave to react for 10 min
- measure activity
- cool to 25 °C
- connect reaction vial with AcOH vial (10 mL) (+vent)
- turn on DH2
- add N₂ to react vial - transfer reaction mixture - remove N₂
- add 5 mL AcOH to reaction vial, shake and transfer with N₂ to the vial with AcOH
- add inlet d to AcOH vial and line e to ACN 15% in water
- (+vent +radioactivity sensor on vial after C18)
- (computer) start command 2A (CE2 d to f) X 4 times
- start command 2B
- start command 2A (again)
- remove needle from C18
- add QMA+2 Al-N (SEP-PAK) (+needle)
- move to elution vial (+vent +rad sensor on first Al-N)
- add d and 2nd N₂ line to Et₂O
- (computer) turn on DH-2
- mass flow 1 mL/min

conc2	}	(x2 times)
d-fill reag loop		
absolute		
1000		
3000 - OK		
conc2		
a-waste		
dispense		
1000		
9999 - OK		

conc2

d-fill reag loop

absolute

1000

3000 - OK

conc2	}	(repeat until shows activity)
f		
dispense		
100		
1000		

- move needle to next vial (+radioactivity sensor on last AI-N)

- (computer)	conc2	}	(repeat until empty activity)
	f		
	dispense		
	100		
	1000		

- measure activity (+vent)

- add N₂

- long needle (pump3) in ACN

- (computer)	pump3
	f-reag line
	aspirate
	100
	1000 - OK

take needle out

pump3
f-reag line
aspirate
5 (air)
1000 - OK

insert needle in the reaction vial

pump3
f-reag line
aspirate
10
1000 - OK

take needle out, insert in MS vial for HPLC

pump3

f-reag line

dispense

35

1000 - OK

- take out sample for HPLC

- put back reaction vial on stirring at 40 °C

- turn DH-1 off

- turn mass flow to 2 mL/min

- dry

↓ (while drying)

- put long needle (pump3) in DMSO vial (+vent)

- (*computer*) pump3

f-reag line

aspirate

400

2000 - OK

pump3

a-waste

dispense

400

2000 - OK

pump3

f-reag line

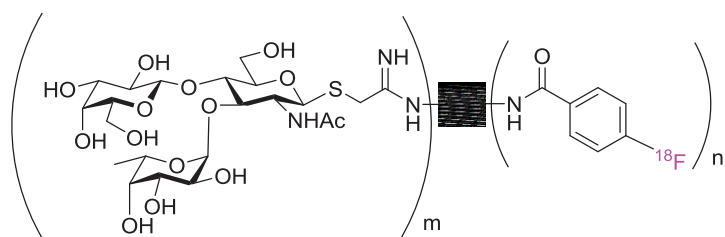
aspirate

400

2000 - OK

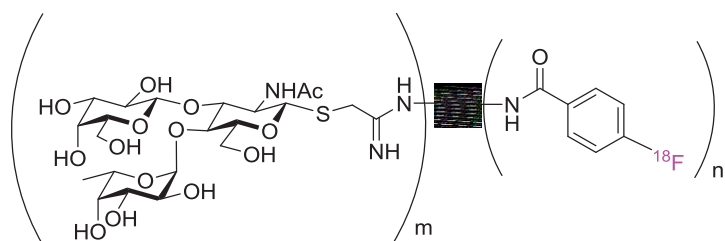
- take out needle and turn upside-down
- aspirate 200
- dispense 200
- dispense 20 (rate 1000)
- turn temperature to 25 °C
- dispense 20 in react vial
- stir for few seconds
- aspirate 20 (rate 500)
- tilt vial and aspirate other 5 (repeat until all is withdrawn)
- move needle to final vial
- dispense 'x' µL (all volume or part of it if multiple vials)

[¹⁸F]-N-succinimidyl-4-F-benzoyl-Le^x-NanoLink 4.24



In the hot cell, to **4.17** (0.2 mg), previously washed with milliQ water and recovered with phosphate buffer (10 mM, pH 7.4, 20 μ L), was added **4.23** in DMSO (15 μ L), and the reaction was stirred for 30 min at 30 °C. The particles were separated by magnet and washed with DMSO (2 x 100 μ L) and buffer (2 x 100 μ L) to afford the title compound **4.24** (0.4 MBq).

[¹⁸F]-N-succinimidyl-4-F-benzoyl-Le^a-NanoLink 4.25



In the hot cell, to **4.19** (0.2 mg), previously washed with milliQ water and recovered with phosphate buffer (10 mM, pH 7.4, 20 μ L), was added **4.23** in DMSO (15 μ L), and the reaction was stirred for 30 min at 30 °C. The particles were separated by magnet and washed with DMSO (2 x 100 μ L) and buffer (2 x 100 μ L) to afford the title compound **4.25** (0.5 MBq).

Nanocapsules for targeted delivery of radioactivity

5.1 Introduction

5.1.1 Radiotherapy

Among all the therapeutic approaches currently used in cancer treatment, radiotherapy is particularly effective in controlling or killing malignant cells. In systemic radioisotope therapy,^[1] an unsealed source of radiation accumulates in the target site as a consequence of the isotope properties (i.e. radioiodine in the thyroid gland) or after the coupling with a suitable targeting agent.^[2,3] Ionizing radiations damage the DNA of the irradiated tissue leading to cellular death. At present, the application of radiotherapy is limited to localised tumours and consequently is usually combined with surgery and chemotherapy.^[1] Radiotherapy presents several drawbacks, such as the risk of radiation-induced cancers and other serious side-effects.^[2] The targeted delivery of anticancer agents^[4] plays a key role against the non-specific action of radiotherapy and its associated adverse effects. Indeed, the effectiveness of a treatment is not only related to its ability to kill cancer cells, but also to the level of damage caused to healthy cells. The degree of change in the patient's quality of life and eventual life expectancy depends on it. Active targeting requires a specific marker (a target)

to be expressed exclusively on disease-associated cells (or at a much higher level than on normal cells) and a suitable targeting agent. The targeting agent will be conjugated to the drug or to a vector loaded with the drug, which will accumulate the active agent at the target site (or trigger a receptor-mediated response). One of the greatest challenges in this approach is defining the optimal targeting agent to selectively and successfully transport the therapeutic agent to the target. These targeting agents may be antibodies for specific antigens,^[5] or ligands for receptors or transporters,^[6] and may be covalently conjugated to the drug *via* an appropriate chemical bond. Drug delivery systems can be divided in two main categories: the first group consists of capsule-like vehicles, including micelles and liposomes.^[7] These systems have the ability to contain a relatively large amount of drug within their self-assembled structures. The second one exploits the covalent attachment of the drugs to the carriers.^[8,9] A diverse range of materials have been used as carriers in this category, including linear polymers and dendrimers^[10] as well as nanomaterials such as nanotubes,^[11] nanorods,^[12] and nanoparticles.^[13]

5.1.2 Carbon nanostructures

Since the advent of the buckminsterfullerenes family in 1980s,^[14] many nanotechnology interests began to gravitate towards carbon-based structures.^[15] Figuratively, some of these structures can be pictured as sheets of rolled-up graphene, in the shape of spheres (fullerenes) or tubules (Fig. 5.1); more precisely single-walled carbon nanotubes come from a single sheet, or multi-walled carbon nanotubes from multiple concentric layers.

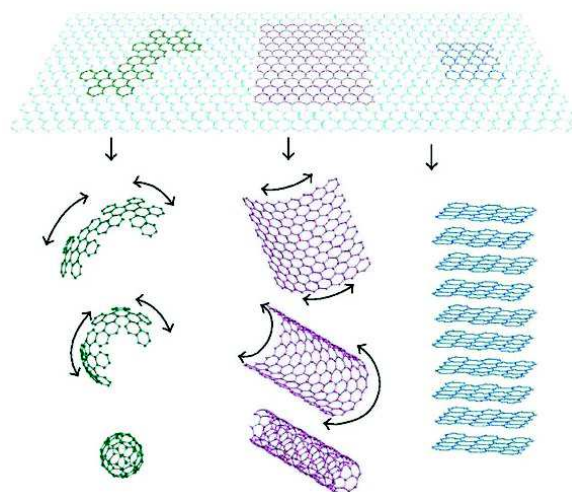


Figure 5.1 Carbon nanostructures.^[16]

This definition is conceptually easy to depict but practically wrong with respect to the mechanism of their formation by arc discharge,^[17] carbon vapour deposition,^[18] plasma torch^[19] or laser ablation.^[20] Even though their structure was already suggested in 1970,^[21] C₆₀ fullerenes were only discovered in 1985,^[14] preceding the multi-walled carbon nanotubes (MWNTs)^[22] in 1991, and single-walled carbon nanotubes (SWNTs)^[23] two years later. Thanks to their unique optical and physical properties,^[24-27] these materials readily found applications in mechanical and engineering fields. The key to their properties can be found in their hyperconjugated sp² bonds. As new materials emerged in the scientific panorama, scientists of different backgrounds started to explore their potential in chemistry, biology and medicine.^[28,29] Carbon nanotubes (CNTs) have been successfully used in neuronal tissue regeneration,^[30] bone cell proliferation^[31] and spinal cord repair,^[32] but also acquired an increasing popularity for bioimaging,^[33] biosensing^[34] and as drug delivery systems.^[35-36] As produced, CNTs are insoluble in almost any aqueous solution and organic solvent; unmodified CNTs can be highly toxic to mammalian cells,^[37] which represents a

great limitation to their application in living systems.^[38-41] Recent strategies to functionalise their side-walls with bioactive molecules have resulted in the generation of stable and biocompatible dispersions. Moreover, supramolecular (non-covalent) binding of aromatic molecules can be easily achieved by π - π stacking of aromatic molecules onto the surface of nanotubes.^[42] CNTs have ultrahigh surface area that permits efficient loading of multiple molecules along the length of the nanotubes such as surfactants,^[43,44] block-copolymers,^[45] peptides,^[46] nucleic acids,^[47] or carbohydrates^[48] and more. In the design of bio-inspired structures, particular attention has been devoted to carbohydrate derivatives, such as synthetic glycopolymers,^[49,50] glycodendrimers,^[51] and glycolipids.^[52]

Previously in our group, nanotubes have been used as a platform for the targeted delivery of radionuclides in the work described by Hong et al. in 2010^[11](Fig. 5.2). They filled steam-purified and shortened^[53] single-walled carbon nanotubes (SWNTs) with Na¹²⁵I^[54] by molten phase capillary wetting. A biantennary carbohydrate (GlcNAc) moiety and a tagging agent (triiodobenzaldehyde),^[55] to detect the functionalisation by electron microscopy, were conjugated by azomethine ylid 1,3-dipolar cycloaddition, also known as the Prato reaction.^[56]

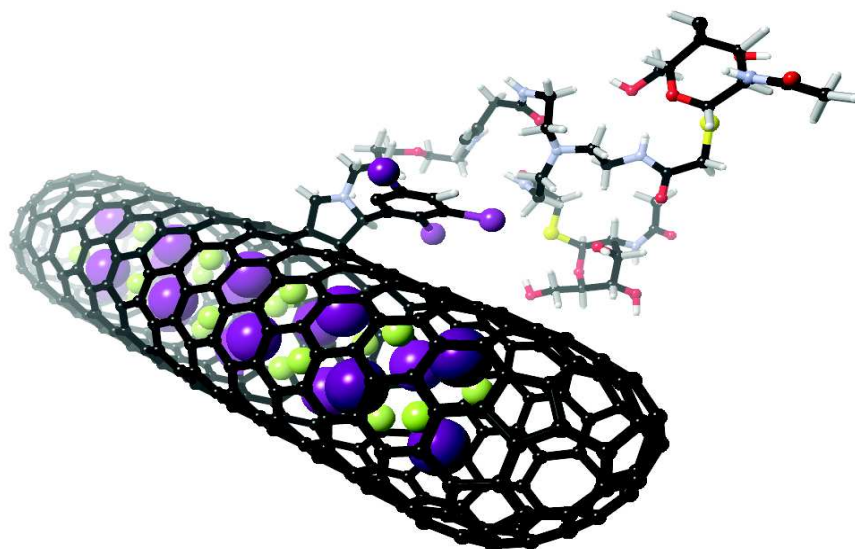


Figure 5.2 SWNTs filled with NaI and functionalised to present a biantennary GlcNAcylated functionality.

The novelty of this system is its capability to enclose the drug (in this case a radionuclide) in the hollow space inside the nanotubes, delivering an unprecedented radiodose density to a specific target. The GlcNAc^[57,58] proved to improve the water dispersibility and biocompatibility^[59] of the SWNTs as well as to redirect ¹²⁵I from the thyroid^[60] to the lungs: their biodistribution was followed after administration in model animals by single-photon emission computed tomography (SPECT) and gamma counting.^[61] We envisage that from this platform, tumor-specific therapeutics and diagnostics can be developed. These nanocapsules represent a sealed source of radiation, with all the advantages of a drug delivery system, capable of being effective not only on localised sites but also on systemic targets. At this point, it appears evident that the next step would be the choice of an appropriate targeting agent tackling specific pathological tissues or processes. Therefore, we decided to take this technology forward by using a different sugar as targeting agent. Galactose was selected to prove that the nanocapsules could selectively accumulate in a different organ (liver^[62]). With

this proof-of-principle experiment we would like to prove that radio-filled nanotubes could serve as platform to study the distribution of the sugar of our Glycomap on a different scaffold. In a wider perspective, we hope this system will be adopted for imaging and therapeutic purposes.

5.2 Synthesis of glyconanotubes

5.2.1 Purification and filling of single-walled carbon nanotubes

Under the supervision of Dr. Gerard Tobias at the Institut de Ciència de Materials de Barcelona (ICMAB) we purified and filled SWNTs. SWNTs synthesised by chemical vapour deposition (CVD),^[63] were purchased from Thomas Swan & Co. Ltd. The material contained amorphous carbon and graphitic particles with an iron core (used as catalyst during the synthesis)^[64] that could prevent a proper functionalisation of the sidewalls of the tubes^[65] and cause cytotoxicity.^[66-68] In order to get rid of such impurities, the SWNTs were purified prior to functionalisation by steam^[53,69] and refluxing in HCl. During the steaming process a lot of carbon was consumed from both the amorphous layer constituting the graphitic shell of the iron particles and by shortening the SWNTs.^[53,69,70] After the purification, the SWNTs were analysed by thermogravimetric analysis (TGA) and transmission electron microscopy (TEM).^[71] The steaming process removed most of the contaminants, but a small amount of graphitic shells survived the treatment and could be seen by TEM (Fig. 5.3) and quantified by TGA.

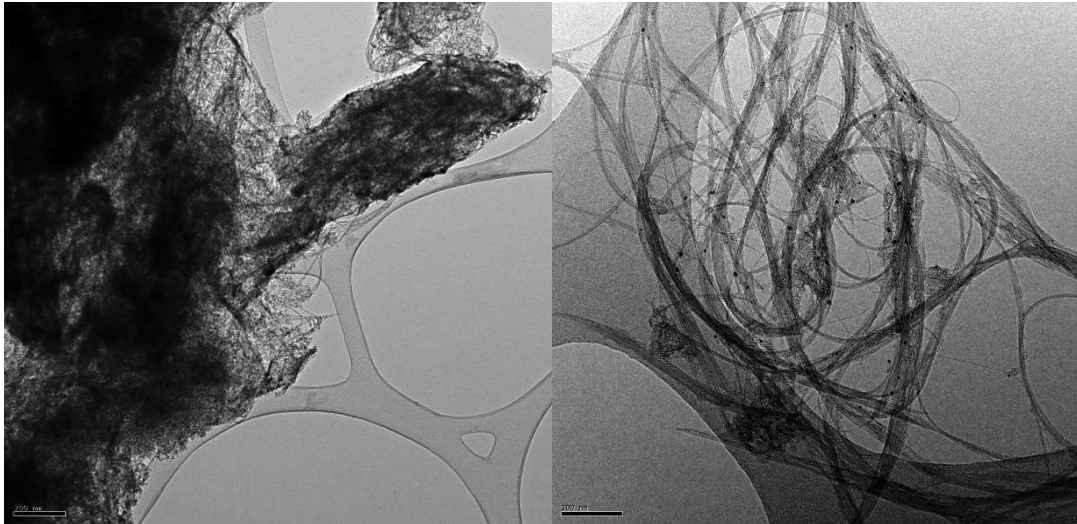
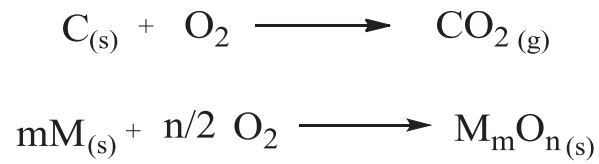


Figure 5.3 TEM images of SWNTs as received (left) and after the purification by steaming (right). Residual Fe particles are visible as black dots.

During the TGA the carbon was consumed and the iron oxidized, leaving a residue of Fe_2O_3 that can be quantified by the following equation: ^[70]



$$M_1(\text{wt } \%) = \frac{m \times R_1 \times \text{MW}_M}{\text{MW}_{\text{M}_m\text{O}_n}}$$

Where m is the stoichiometric constant, and MW_M and $\text{MW}_{\text{M}_m\text{O}_n}$ the molecular weights for the metal and the metal oxide, respectively.

The purified and shortened SWNTs were filled by molten-phase capillary wetting^[72] (Fig. 5.4).

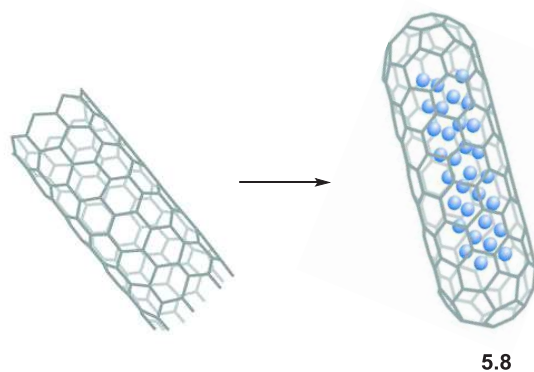


Figure 5.4 Filling of SWNTs with NaI. The ends of the tubes close spontaneously during the cooling process.

After filling, the temperature at which the NaI containing SWNTs (NaI@SWNTs) burn changes dramatically (Fig. 5.5).

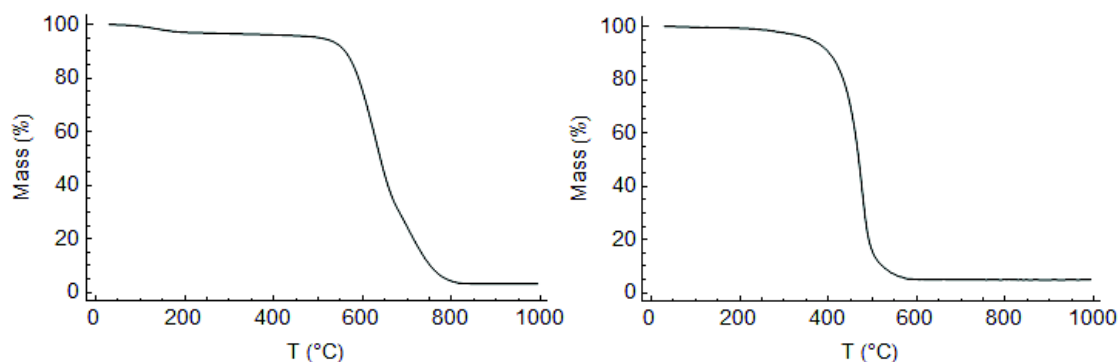


Figure 5.5 TGA of SWNTs before (left) and after filling (right) with NaI. The temperature at which the NTs burn is lower after the filling.

The filling yield can be calculated by the following equation:^[70]

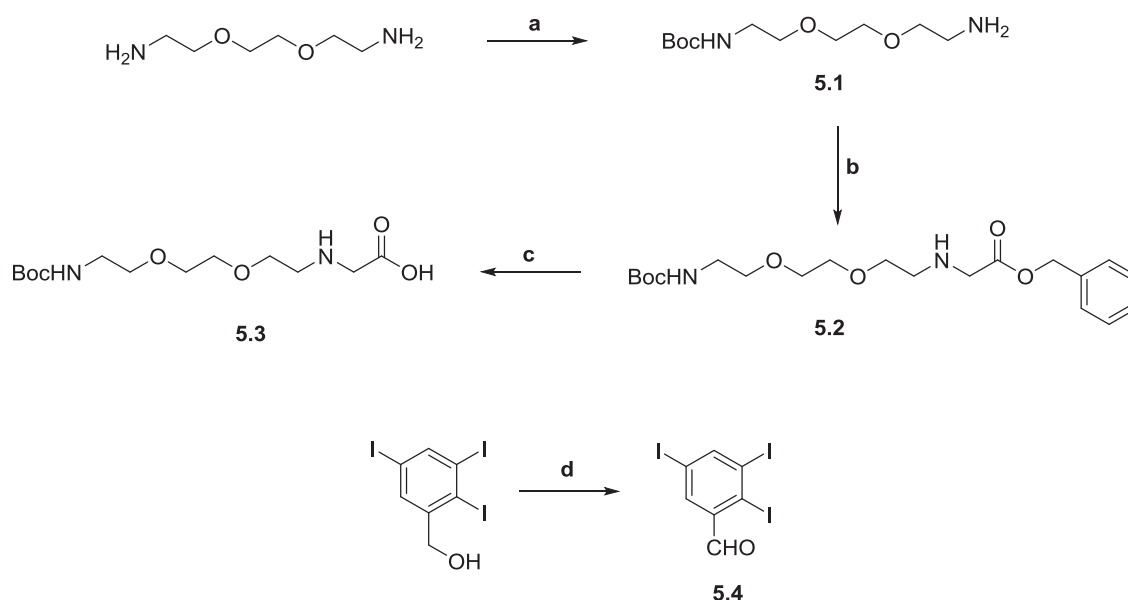
$$FY \text{ (wt \%)} = \frac{100 (R_2 - R_1)}{R_A - R_1}$$

Where R_1 , R_2 , and R_A are the TGA residues in air for the empty carbon nanotubes, the filled carbon nanotubes, and the bulk filling material (A), respectively. The residue observed was 3.27% for the empty SWNTs (left) and 11.93% for NaI@SWNTs.

5.2.2 Functionalisation of NaI@SWNTs

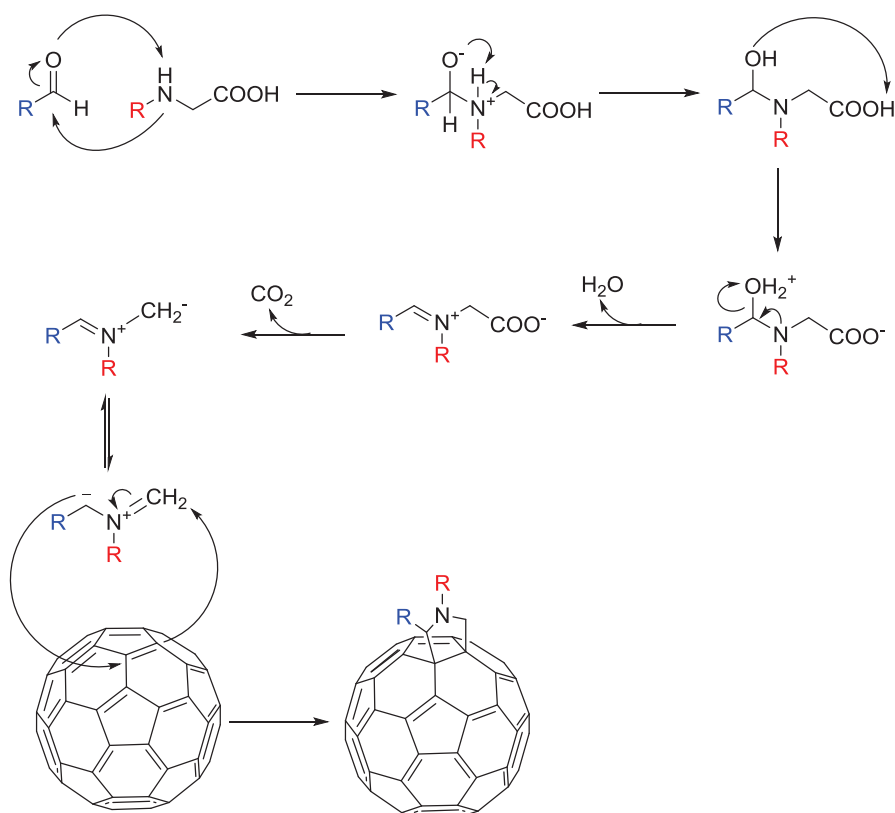
The first modification of the carbon surface consisted of coupling a short linker, which connected the targeting agents to the nanotubes. We improved the original strategy devised by Hong,^[11] designing a convergent approach in which the branched moiety is introduced at the linker level. This represents a more convenient way of preparing a small library of derivatives.

All the reactions for the functionalisation of the carbon surfaces have first been performed on C₆₀-fullerenes. Fullerenes are a good model system to study the reactivity of the nanotubes and can be characterised with common spectroscopic techniques, while small molecules on nanotubes are hardly detectable. We started by using the same linker and the same type of functionalisation published in 2010^[11] following the procedure reported (Scheme 5.1).



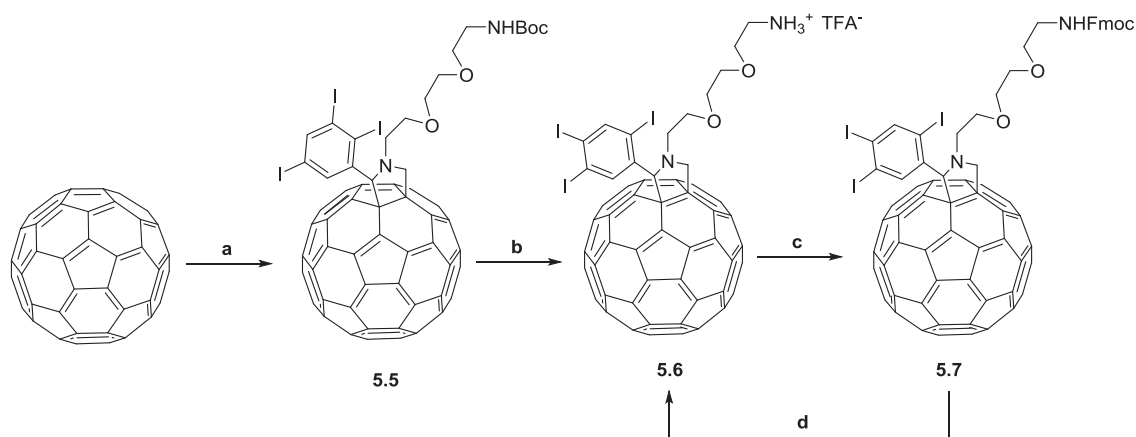
Scheme 5.1 Reagents and conditions: a) di-tert-butyl dicarbonate, 1,4-dioxane, r.t., 95%; b) benzylbromoacetate, 1,4-dioxane, r.t., 98%; c) H₂, Pd/C (10%), MeOH, r.t., 87%; d) PCC, THF, r.t., 53%.

The functionalisation proceeds through the formation of the azomethine ylide which attacks the double bonds on fullerenes (or CNTs) in a 1,3-dipolar cycloaddition^[56] (Scheme 5.2).



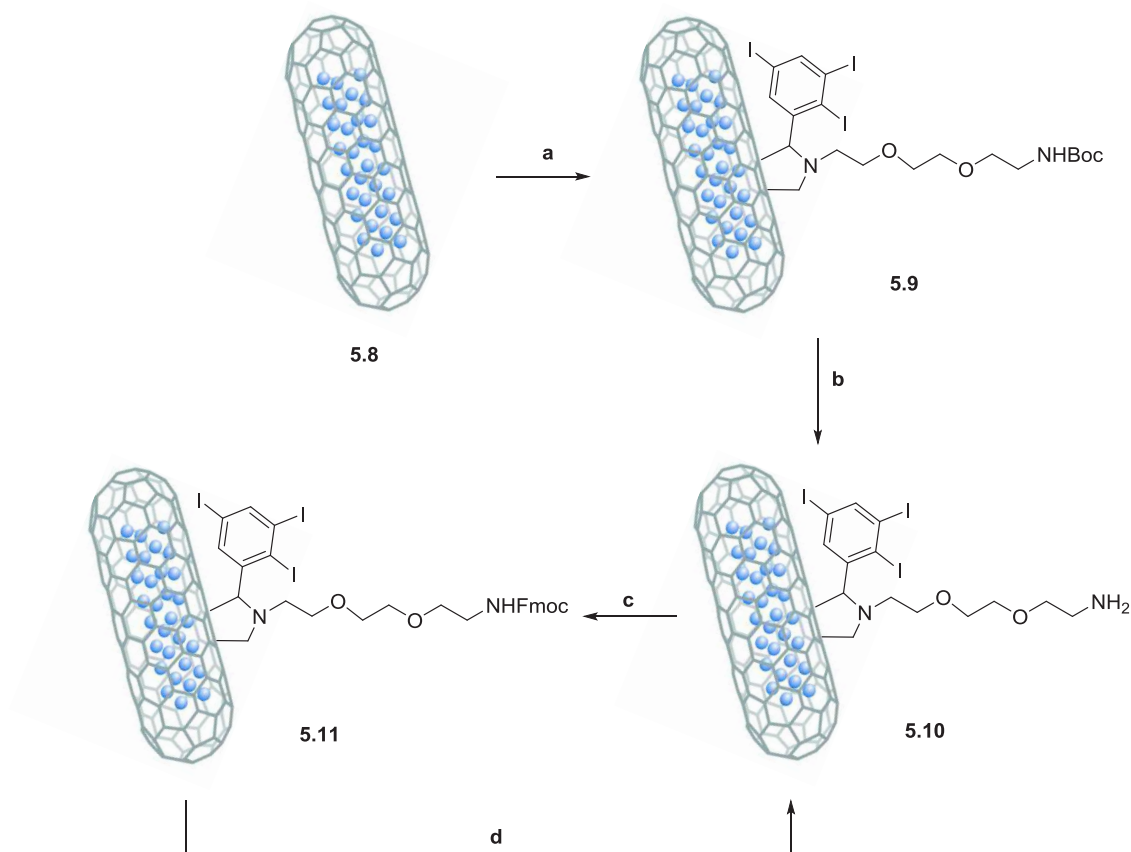
Scheme 5.2 Mechanism of the formation of the azomethine ylide and 1,3-dipolar cycloaddition on the double bond of the fullerene.

Functionalisation has been carried out on C₆₀ fullerenes and characterised by Fmoc numbering and microanalysis before moving onto nanotubes (Scheme 5.3).



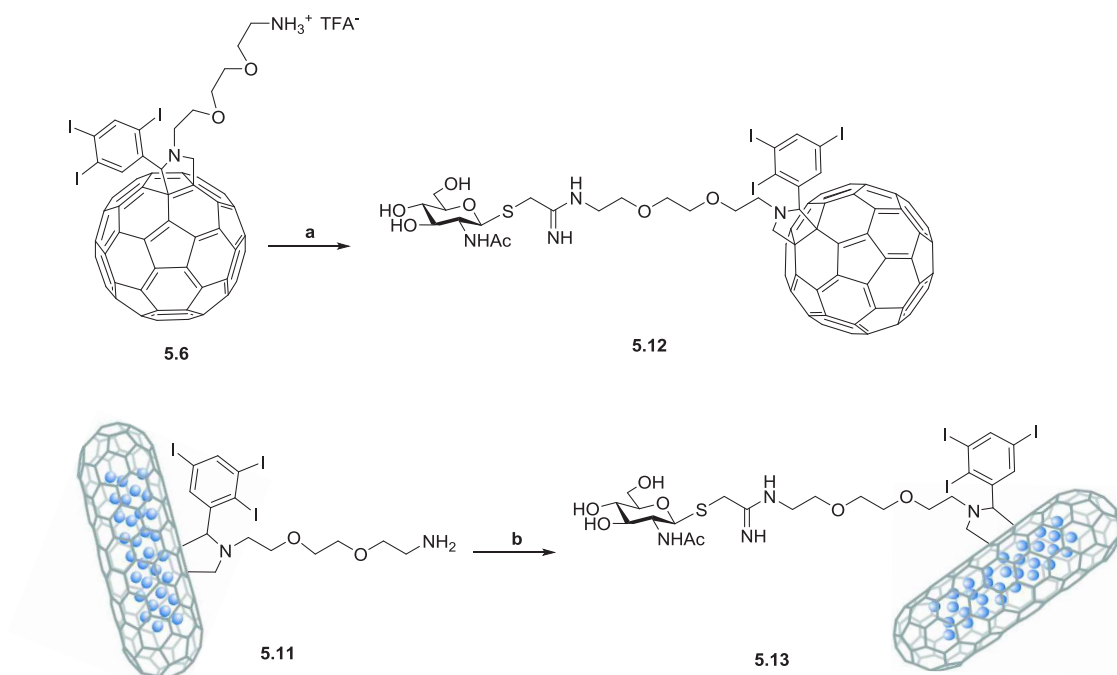
Scheme 5.3 Reagents and conditions: a) **5.3**, **5.4**, toluene, 110 °C, 46%; b) TFA, DCM, r.t., 63%; c) Fmoc chloride, DCM, DIPEA, r.t., 84%; d) DMSO:DMF:DBU (25:24:1), r.t..

Once SWNTs were modified with the linker, a substantial difference was observed. While the procedure adopted was the same (Scheme 5.4), the reaction time required by the Prato reaction^[56] on the CNTs was significantly longer (4 days) than for the fullerenes (2 h). This can be partially attributed to the fact that SWNTs form a dispersion in DMF while fullerenes are reacting in solution, therefore the reaction for the CNTs occurs in the solid/solution interphase. Fullerenes are also known to be more reactive than CNTs due to their more pronounced curvature, which increases the strain (and therefore the reactivity) on the double bonds.



Scheme 5.4 Reagents and conditions: a) **5.3**, **5.4**, DMF, 130 °C; b) TFA, DCM, r.t.; c) Fmoc chloride, DCM, DIPEA, r.t.; d) DMSO:DMF:DBU (25:24:1), r.t..

Functionalised SWNTs showed higher dispersibility in some organic solvents, though it would not be correct to confuse the ability of any functionalised SWNTs to form stable suspensions with solubility. While Hong and co-workers^[11] pre-formed the biantennary carbohydrate units before coupling them to the primary amine of the linker we decided to use GlcNAc-IME reagent for a direct conjugation (Scheme 5.5).

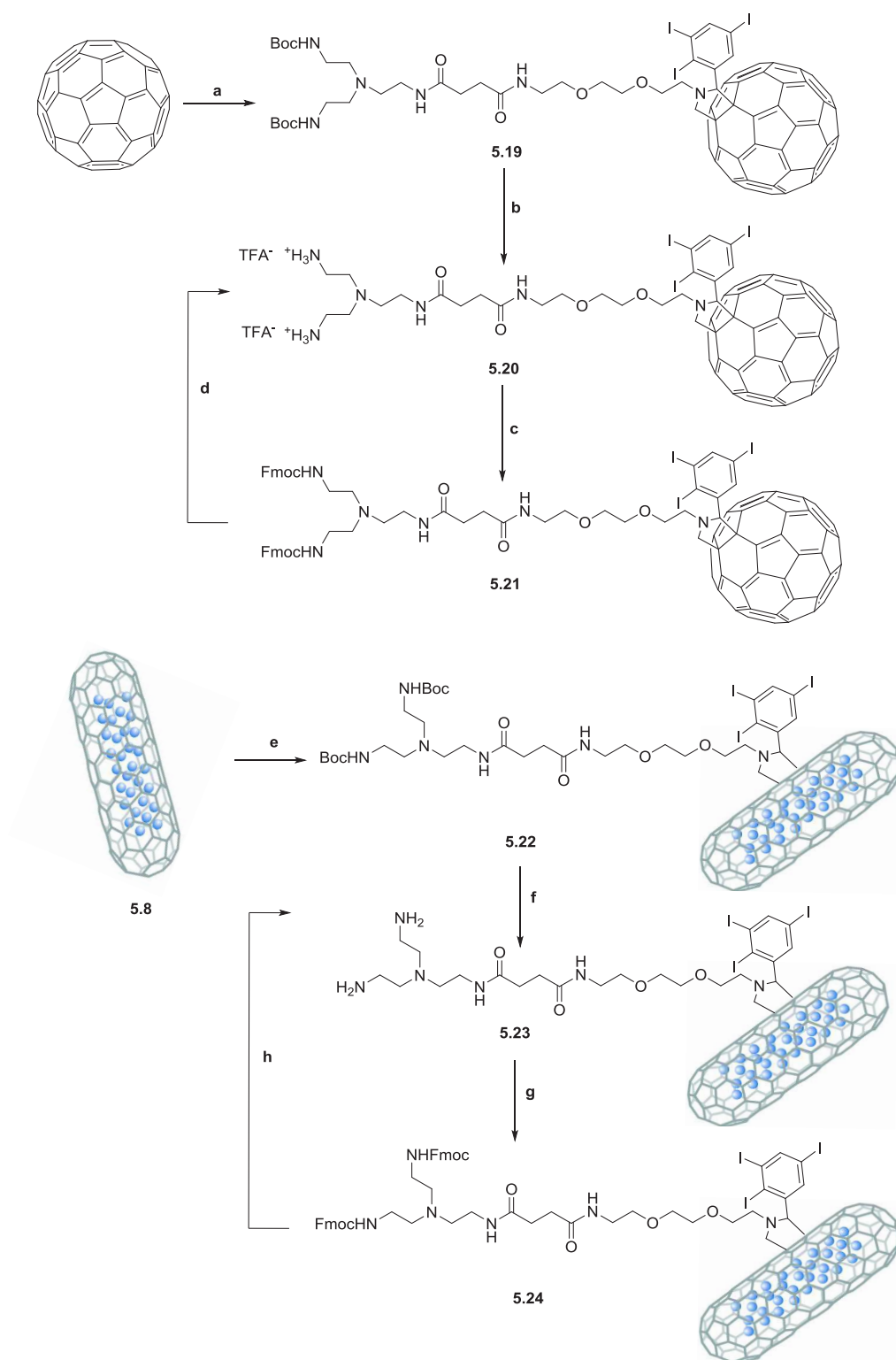


Scheme 5.5 Reagents and conditions: a) **2.80**, MeOH:DMSO (2:1), r.t., 64%;
b) **2.80**, MeOH:DMSO (2:1), r.t..

The mono-glycosylated linker appended on both fullerenes (1 copy) and SWNTs (multiple copies) has been fully characterised and later compared to the biantennary analogue.

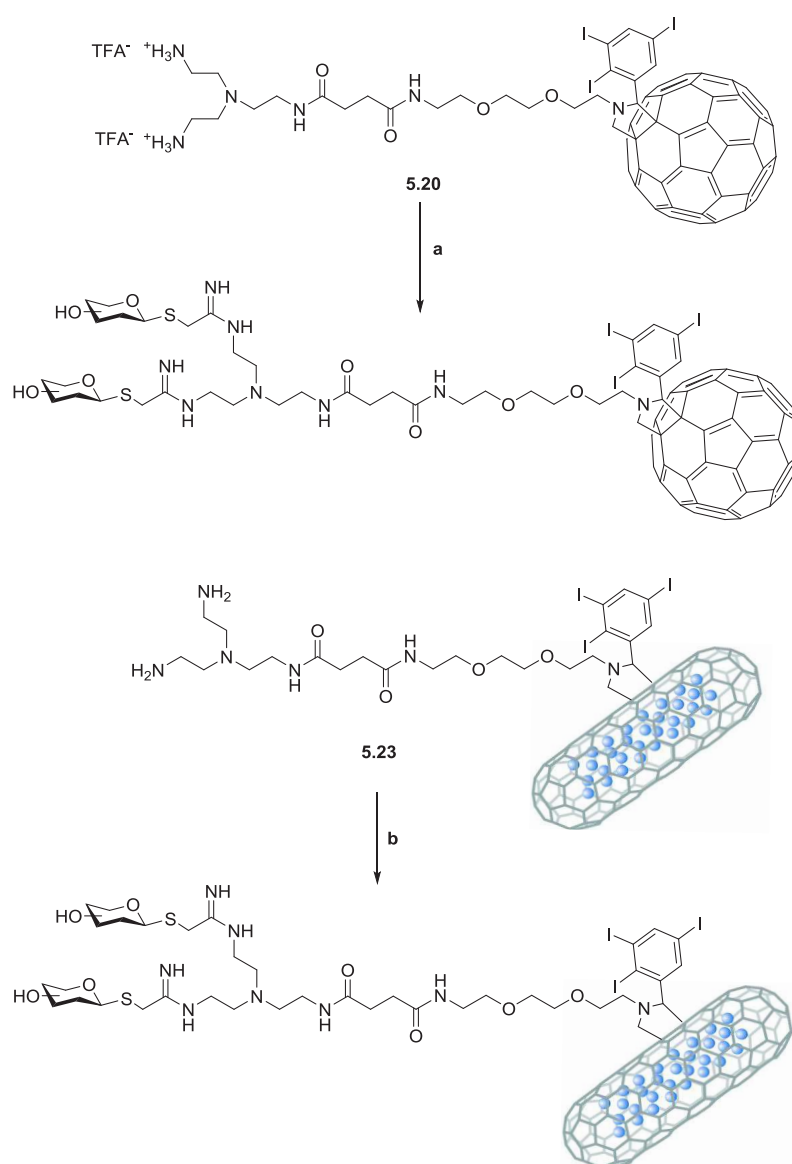
The biantennary linker was synthesised to match exactly the same linker presented by Hong et al.,^[11] up to the sugar unit. The synthetic route was itself heavily inspired by their approach, combining the branching section to the linear spacer (Scheme 5.6). The reason we did not want to make any change to the spacer was to not introduce a new variable once comparing the two glycosylated systems.

Linker **5.18** was used immediately to functionalise C₆₀ fullerenes and NaI@SWNTs (Scheme 5.7).



Scheme 5.7 Reagents and conditions: a) **5.18**, **5.4**, toluene, 110 °C, 54%; b) TFA, DCM, r.t., quant.; c) Fmoc chloride, DCM:DMF (1:1.5), DIPEA, r.t., 85%; d) DMSO:DMF:DBU (25:24:1), r.t.; e) **5.18**, **5.4**, DMF, 130 °C; f) TFA, DCM, r.t.; g) Fmoc chloride, DCM, DIPEA, r.t.; h) DMSO:DMF:DBU (25:24:1), r.t..

Finally, the last point to address was the glycosylation of functionalised fullerenes and nanotubes with GlcNAc-IME and Gal-IME (Scheme 5.8). The reason behind changing the carbohydrate moiety to galactose was eventually to redirect the radio-filled (hot) nanotubes to the liver,^[62] proving we would be capable of targeting this huge (in physiological terms) drug delivery system by simple manipulation of the carbohydrates, regardless of the properties of the nanotubes which will be relegated to playing the role of simple cages.

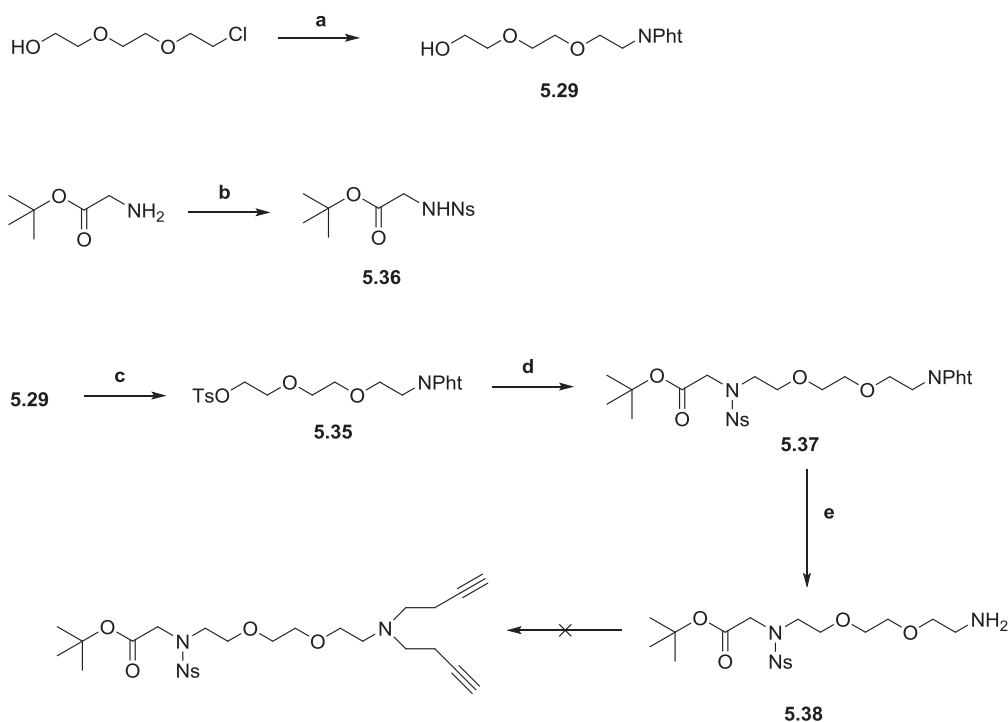


Scheme 5.8 Reagents and conditions: a) **2.80** or **2.77**, MeOH:DMSO (2:1), r.t., 54%; b) **2.80** or **2.77**, MeOH:DMSO:DCM (2:1:0.25), r.t..

Fmoc numbering revealed a degree of functionalisation of 0.482 mmol/g for the mono-glycosylated fullerenes while the bifunctional equivalent reached about twice that value (0.742 mmol/g). Obtaining double the functionalisation for the second fullerene construct is not surprising, considering that it is already fully characterised by MS and NMR spectrometry. However, this is not the case for the nanotubes as their characterisation by microanalysis is atypical. The Fmoc numbering gave a functionalisation of 0.162 mmol/g for mono-glycosylated tubes, in line with the reported value of 0.17 mmol/g.^[11] Di-glycosylated nanotubes, similarly to fullerenes, also reported twice the functionalisation of the cognate mono-derivative (0.34 mmol/g). These results have been compared to a second round of data obtained with a different analytical technique. Elemental analysis recorded an approximate content of 67% of C, 1% of H and 3% of N for all glycosylated nanotubes. Values did not vary significantly for any given sample, therefore no conclusions can be drawn for these results. Knowing the oxygen content might have been more helpful, but unfortunately this could not be measured. No significant difference was observed either between nanotubes functionalised with mono- and divalent linker (without sugars), as the values of the elements detected were still oscillating around the same numbers as per the glycosylated CNTs. The only appreciable information that the elemental analysis revealed was the appearance of H and N content for functionalised samples with respect to filled unmodified CNTs.

5.2.3 Design of a new linker

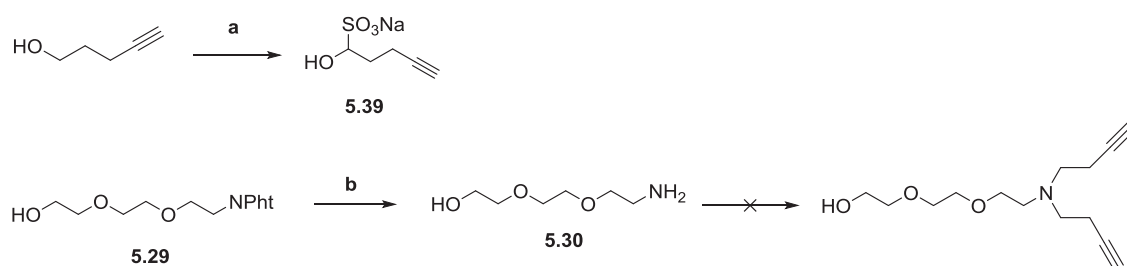
Following the successful glycosylation of the amines of the new biantennary linker we designed, we conceived a new spacer which would let us couple our sugars on the nanotubes using a different type of chemistry. Copper catalysed azide-alkyne cycloaddition (CuAAC) has been used before to decorate CNTs^[73] thanks to its efficiency and practicality. Mindful of the GlcNAc-azido derivative synthesised previously for protein modification (Chapter 3.3.3), we decided to install a terminal alkyne on each end of our bifunctional spacer. The synthesis of this linker had to be redesigned and during this process we encountered not a few difficulties (Scheme 5.9).



Scheme 5.9 Reagents and conditions: a) potassium phthalimide, DMF, 100 °C, 84%; b) nosyl chloride, collidine, DCM, r.t., 82%; c) tosyl chloride, pyridine, DCM, r.t., 62%; d) **5.36**, ACN, Cs₂CO₃, 85 °C, 25%; e) hydrazine, EtOH, 85 °C, quant..

First, we tried to synthesise this new di-alkyne linker by installing a phthalimide-masked amine^[74] on 2-[2-(2-chloroethoxy)ethoxy]ethanol and activate the alcohol at the other end to couple it with protected ^tBu-Gly. The amine of the amino acid has been protected with nosyl group^[75] to prevent dialkylation, constituting an orthogonal protecting group with respect to the phthalimide. After this conjugation the dialkylation of the deprotected amine was attempted under several conditions (heating, several microwaves (MW) cycles, excess of reagent in multiple additions, different solvents) but unfortunately no product was observed.

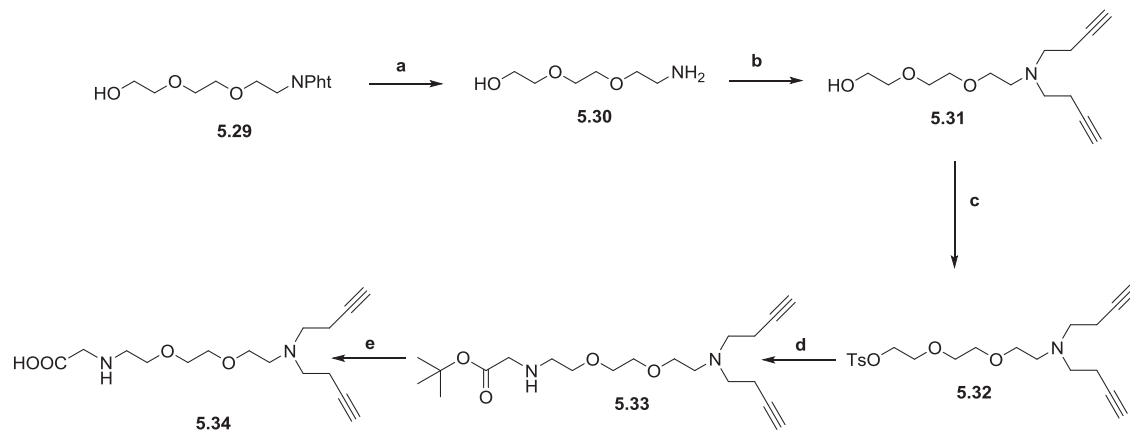
This route has therefore been abandoned but from the starting material **5.29** we tried a different approach (Scheme 5.10).



Scheme 5.10 Reagents and conditions: a) i. PCC, DCM, EtOH, r.t.; ii. NaHSO₃, EtOH, H₂O, r.t., 87%; b) hydrazine monohydrate, EtOH, 100 °C, quant..

On this occasion we attempted directly the dialkylation of the deprotected primary amine **5.30**^[76] with the salt of the hydrated form of 4-pentynal **5.39** (equivalent to the aldehyde itself). Unfortunately even here, after trying harsher conditions, no product was formed.

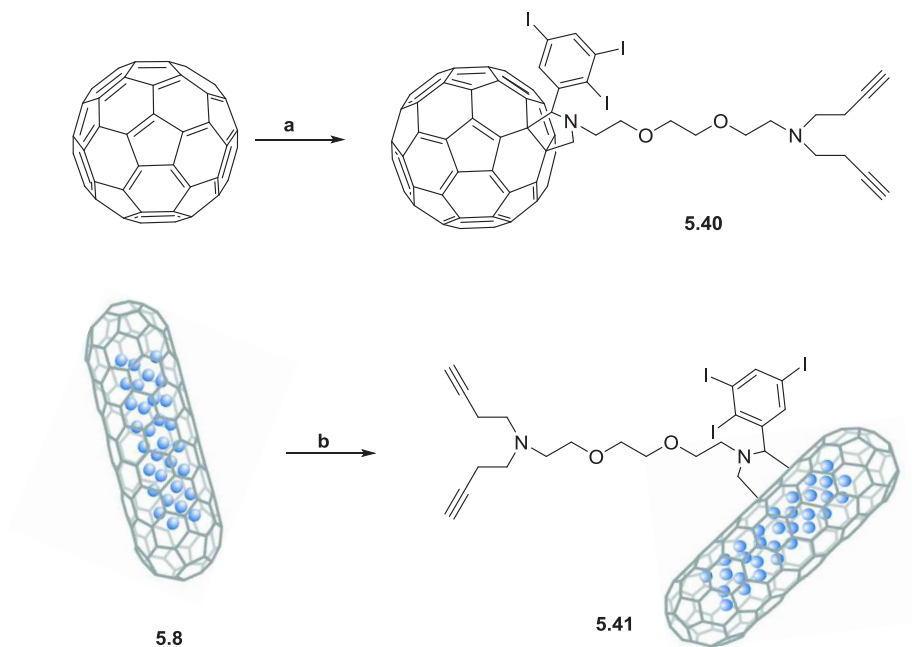
Finally we found a way around the hardship of installing the two alkyne moieties (Scheme 5.11).



Scheme 5.11 Reagents and conditions: a) hydrazine monohydrate, EtOH, 100 °C, quant.; b) KI, Cs₂CO₃, 4-bromobutyne, ACN, MW 85 °C, 54%; c) tosyl chloride, pyridine, DCM, r.t., 51%; d) ^tBu-Gly, ACN, Cs₂CO₃, r.t., 8%; e) TFA, DCM, r.t., quant..

Again, from the same precursor **5.29**, we deprotected the amine and forced the dialkylation with multiple additions of 4-bromobutyne and cycles of MW. After turning that corner, the conjugation of the ^tBu-Gly (and various deprotection steps) was straightforward.

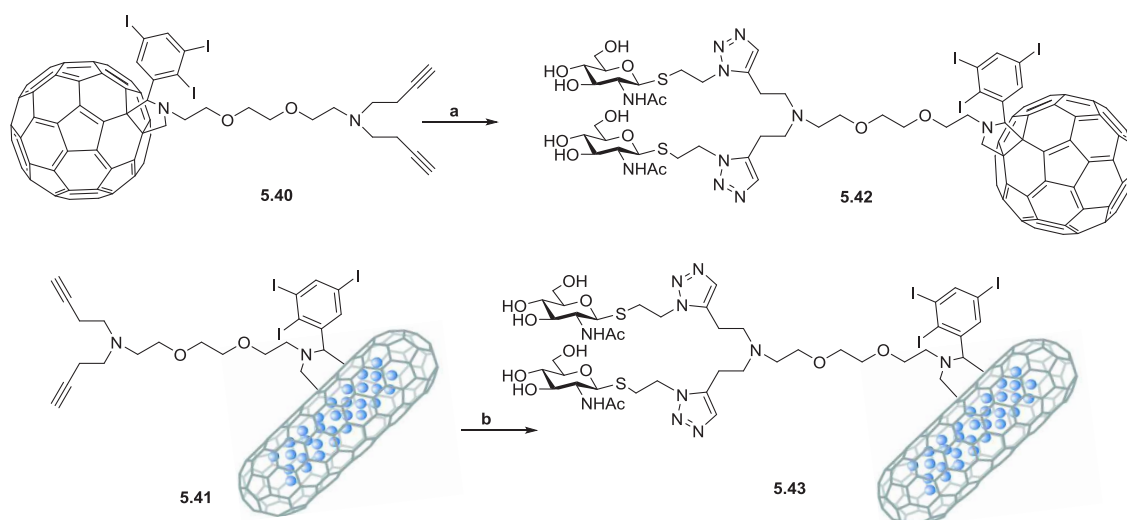
Again, via Prato reaction,^[56] we achieved the functionalisation of both fullerenes and NaI@SWNTs (Scheme 5.12).



Scheme 5.12 Reagents and conditions: a) **5.34**, **5.4**, toluene, 120 °C, 23%; b) **5.34**, **5.4**, DMF, 130°C.

Although with this new linker the Fmoc numbering was no longer possible, successful functionalisation was still confirmed by elemental analysis.

GlcNAc-N₃ was finally coupled on the new linker by CuAAC (Scheme 5.13).



Scheme 5.13 Reagents and conditions: a) **3.7**, Cu(I)·TTMA, MeOH:DMSO (2:1), r.t., 87%; b) **3.7**, Cu(I)·TTMA, MeOH:DMSO (2:1), r.t..

As final consideration, we succeeded in our intent to explore the variability of this carbon-based platform both at the linker and at the sugar level. The key to understanding whether we have been successful also in the perspective of its applications will be the *in vivo* experiments with the hot-filled nanotubes, of which the synthesis is reported below.

5.3 Characterisation of glyconanotubes by electron microscopy

While small molecules are difficult to detect on CNTs, other techniques can tell us more about their properties. Electron microscopy (EM) is a powerful technique based on the interactions of electrons with the sample. A beam of electrons is transmitted through the specimen and the deviation of the electrons collected and

elaborated to form an image. Transmission electron microscopy (TEM) differs from scanning electron microscopy (SEM) as in the latter the electron beam does not go through the sample but is scanned across it. The two techniques can be combined in the STEM mode.

5.8 and **5.9** were characterised by TEM and elemental analysis (Chapter 5.2.2). In HAADF-STEM (High-Angle Annular Dark-Field Scanning Transmission Electron Microscopy) mode the intensity of the signal is proportional to the atomic number. In the TEM image (Fig. 5.6) of **5.9** (right) we can observe bright dots on the bundle areas corresponding to the iodine atoms (indicated by arrows) which do not appear in the image of **5.8** (left).

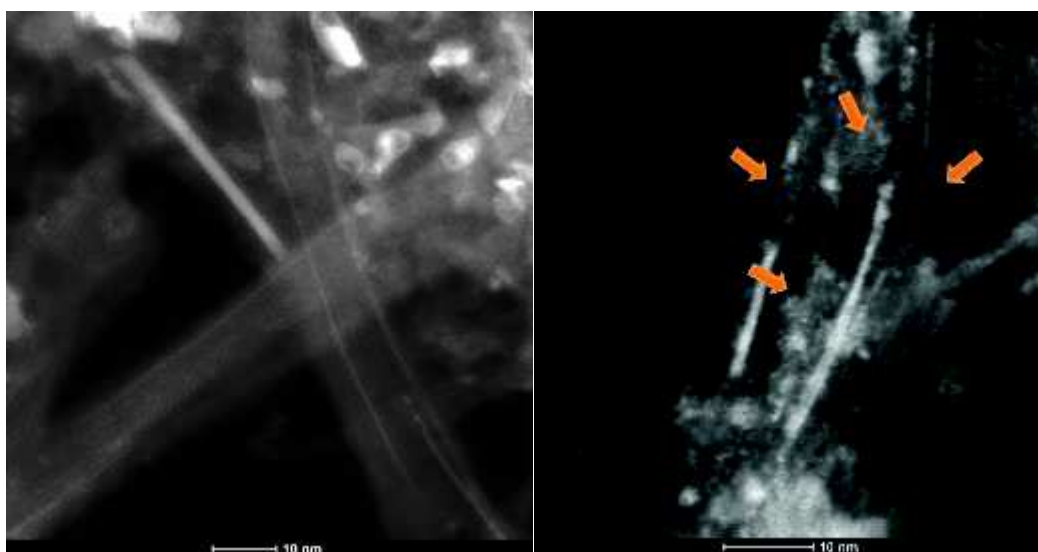


Figure 5.6 TEM of SWNTs before (left) and after functionalisation (right). The arrows show the bright dots corresponding to the iodine atoms introduced with the functionalisation.

The intensity profiles of a cross-section of the samples (Fig. 5.7) show that for the iodine-tagged sample **5.9** (left) the intensity is higher on the bundle area compared to the non-functionalised sample **5.8** (right).

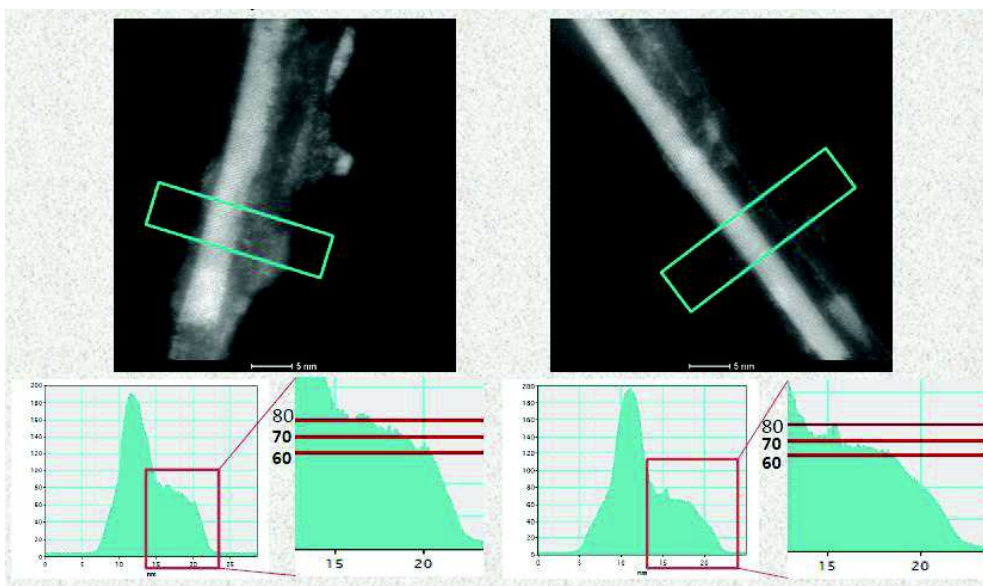


Figure 5.7 TEM of SWNTs before (right) and after functionalisation (left). The iodine content for a section of NTs is considerably higher after functionalisation.

In TEM Bright Field mode combined with EDX (Energy-Dispersive X-ray) analysis (Fig. 5.8) iodine peaks for functionalised **5.9** (right) were seen while no signal was detected for **5.8** (left).

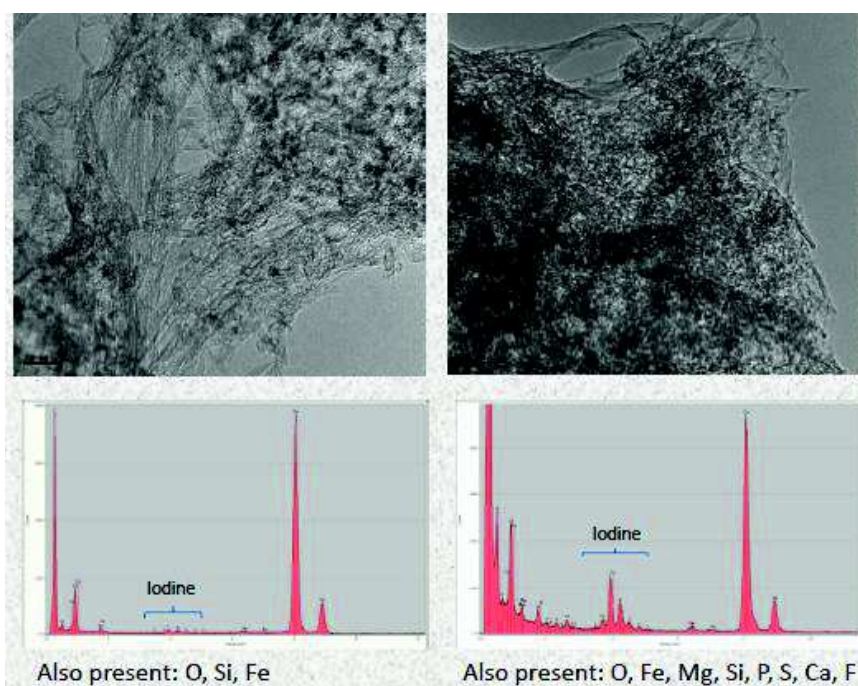
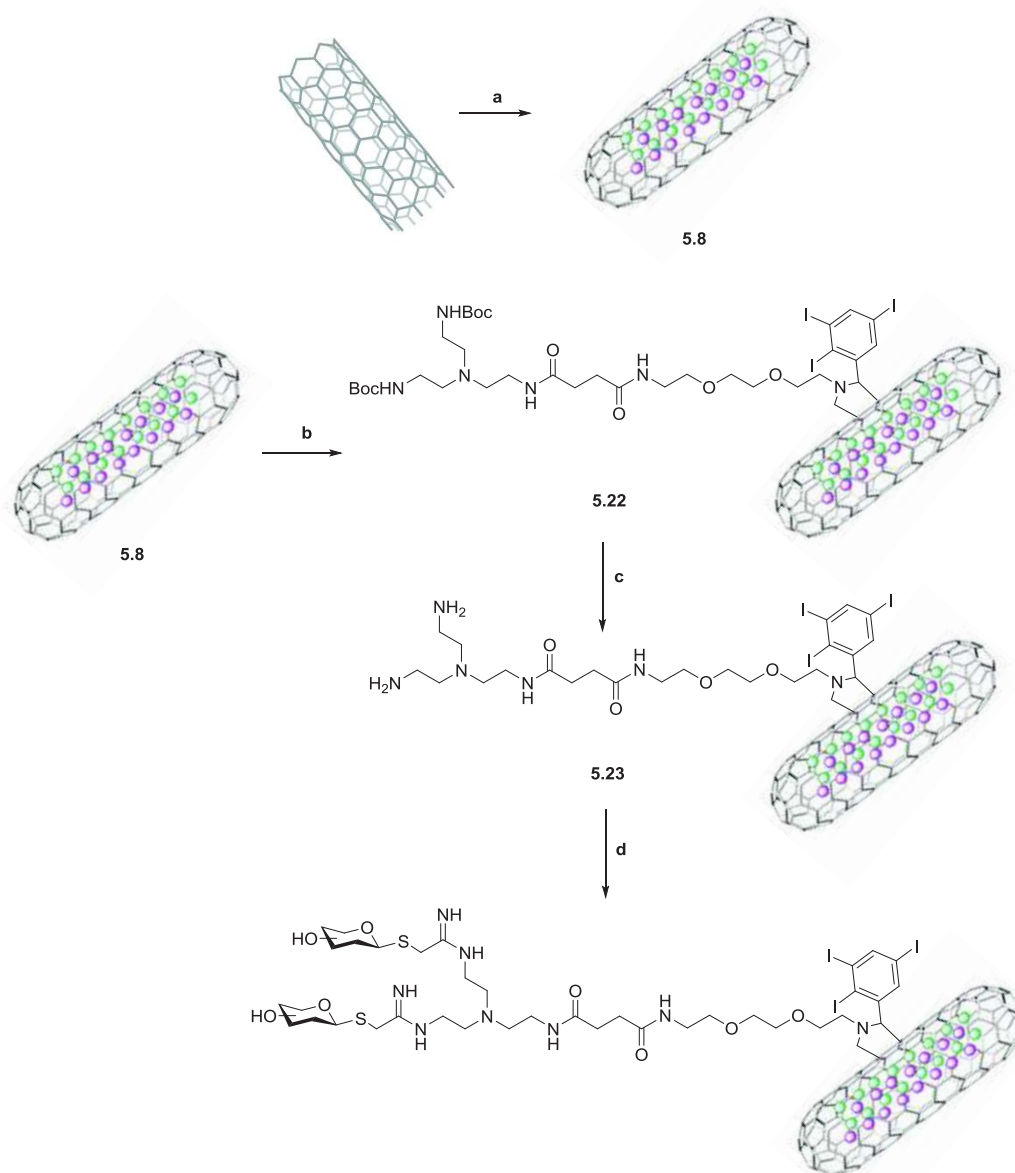


Figure 5.8 TEM of SWNTs before (left) and after filling (right) with NaI. The iodine content is visibly increased from background after filling.

5.4 Synthesis of Na¹²⁵I@SWNTs

Following the general procedure used to fill the SWNTs with ordinary NaI, we proceeded in using Na¹²⁵I (Scheme 5.14).

¹²⁵I is a radioisotope of iodine decaying by emission of gamma and X-rays. It is commercially available in the form of aqueous solution, so a little manipulation is necessary before sealing the quartz ampoule with the empty open-ended nanotubes. After loading with the salt solution and the nanotubes, the vial is left under high vacuum until completely dry, before sealing the tube and heating in the furnace. This step is crucial, as the expansion of any liquid or gas present in the vial would lead to a breakage and a leak of the material.



Scheme 5.14 Reagents and conditions: a) SWNTs, Na^{125}I , $900\text{ }^\circ\text{C}$; b) **5.18**, **5.44**, DMF, $130\text{ }^\circ\text{C}$; c) TFA, DCM, r.t.; d) **2.80**, **5.46**, MeOH:DMSO:DCM (2:1:1), r.t..

Hot nanotubes have been glycosylated with GlcNAc, in order to reproduce the results published before,^[11] and Gal as new targeting agent to redirect the construct to a different organ, as proof-of-principle to test if this system would be a suitable candidate for our Glycomap. If the accumulation observed before was due to the targeting ability of the sugars decorating the outer shell of the nanotubes, with Gal we should observe an accumulation in the liver.^[62]

5.5 Conclusions

In this chapter we expanded our portfolio of glycosylated nanosystems by exploring CNTs. From the idea Hong and co-workers developed a few years ago, we aimed to extend the delivery of the radiodose loaded in the cavity of the nanotubes to different organs, receptors, physiological processes and possibly pathological sites for both imaging and radiotherapy by changing the carbohydrates responsible for their distribution. In order to achieve our aim we had to revise the synthesis of the spacer connecting the side walls of the nanotubes and the sugar units, to translocate the branching at the linker level, prior to glycosylation with our library of sugars. In the process of redesigning this linker we changed the coupling of the sugar from a peptide coupling to an IME-mediated one. We also designed a new linker terminating with alkynes, which can be readily modified with azido sugars *via* CuAAC reaction. We suggest that this new chemistry could encourage other research groups who are familiar with using CuAAC reactions to modify macromolecular structures, to explore the potential of the system we presented here. In future, we will glycosylate this platform with our library of carbohydrates, confident that a new scaffold will contribute to our ambitious aim of unravelling carbohydrate interactions in biological systems.

5.6 References

- [1] G. Gordon, M. J. Peckham, *Int. J. Rad. Oncol. Biol. Phys.* **1979**, 5, 85-91.
- [2] D. Azria, M. Ozsahin, A. Kramar, S. Peters, D. P. Atencio, J. Dubois, N. E. A. Crompton, and B. S. Rosenstein, *Clin. Cancer Res.* **2008**, 14, 6284-89.
- [3] G. E. Hanks, *NCI Monogr.* **1988**, 7, 75-84.
- [4] C. Kim and S. Lim, *Arch. Pharm. Res.* **2002**, 25, 229-39.
- [5] E. Dadachova, A. Nakouzi, R. A. Bryan, and A. Casadevall, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, 100, 10942-47.
- [6] T. M. Allen, *Nature* **2002**, 2, 750-83.
- [7] G. Tiwari, R. Tiwari, B. Sriwastawa, L. Bhati, S. Pandey, P. Pandey, and S. K. Bannerjee, *Int. J. Pharm. Investig.* **2012**, 2, 2-11.
- [8] D. Peer, J. M. Karp, S. Hong, O. C. Farokhzad, R. Margalit, and R. Langer, *Nat. Nanotechnol.* **2007**, 2, 751-60.
- [9] O. M. Koo, I. Rubinstein, and H. Onyuksel, *Nanomedicine* **2005**, 1, 193-212.
- [10] E. R. Gillies, J. M. J. Fréchet, and J. M. J. Fréchet, *Drug Discov. Today* **2005**, 10, 35-43.
- [11] S. Y. Hong, G. Tobias, K. T. Al-Jamal, B. Ballesteros, H. Ali-Boucetta, S. Lozano-Perez, P. D. Nellist, R. B. Sim, C. Finucane, S. J. Mather, M. L. H. Green, K. Kostarelos, and B. G. Davis, *Nat. Mater.* **2010**, 9, 485-90.
- [12] Y. Xiao, H. Hong, V. Z. Matson, A. Javadi, W. Xu, Y. Yang, Y. Zhang, W. Engle, R. J. Nickles, W. Cai, D. A. Steeber, and S. Gong, *Theranostic* **2012**, 2, 757-68.
- [13] L. Zhang, H. Chen, L. Wang, T. Liu, J. Yeh, G. Lu, and L. Yang, *Nanotechnol. Sci. Appl.* **2010**, 159-70.

- [14] R. E. Kroto, H. W., Heath, J. R. O'Brien, S. C. Curl, R. F. Smalley, *Nature* **1985**, 318, 162-63.
- [15] O. A. Shenderova, V. V Zhirnov, D. W. Brenner, O. A. Shenderova, V. V Zhirnov, and D. W. Brenner, *Crit. Rev. Solid State Mater. Sci.* **2002**, 27, 227-356.
- [16] K. S. Geim, A. K., Novoselov, *Nat. Mater.* **2007**, 6, 183-91.
- [17] N. Arora and N. N. Sharma, *Diam. Relat. Mater.* **2014**, 50, 135-50.
- [18] M. Kumar and Y. Ando, *J. Nanosci. Nanotechnol.* **2010**, 10, 3739-58.
- [19] M. P. J. Pacheco, M. Valdivia, L. Bernal, and R. Valdivia, *Brazilian J. Phys.* **2004**, 34, 1684-88.
- [20] J. Chrzanowska, J. Hoffman, A. Ma, M. Mazurkiewicz, and T. A. Kowalewski, *Phys. Status Solidi B* **2015**, 1867, 1860-67.
- [21] E. Osawa, *Kagaku* **1970**, 25, 854-63 (in Japanese).
- [22] S. Iijima, *Nature* **1991**, 354, 56-58.
- [23] T. Iijima, S., Ichihashi, *Nature* **1993**, 363, 603-05.
- [24] R. J. Chen, S. Bangsaruntip, K. A. Drouvalakis, N. Wong, S. Kam, M. Shim, Y. Li, W. Kim, P. J. Utz, and H. Dai, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, 100, 4984-89.
- [25] N. Wong, S. Kam, T. C. Jessop, P. A. Wender, and H. Dai, *J. Am. Chem. Soc.* **2004**, 126, 6850-51.
- [26] A. Bianco, K. Kostarelos, D. Partidos, and M. Prato, *Chem. Commun.* **2005**, 571-77.
- [27] P. Cherukuri, S. M. Bachilo, S. H. Litovsky, and R. B. Weisman, *J. Am. Chem. Soc.* **2004**, 126, 15638-39.
- [28] Y. Lin, S. Taylor, H. Li, K. A. S. Fernando, L. Qu, W. Wang, L. Gu, B. Zhou,

- Y. Sun, and Y. Lin, *J. Mater. Chem.* **2004**, 14, 527-41.
- [29] D. Tasis, N. Tagmatarchis, A. Bianco, and M. Prato, *Chem. Rev.* **2006**, 106, 1105-36.
- [30] M. P. Mattson, R. C. Haddon, and A. M. Rao, *J. Mol. Neurosci.* **2000**, 14, 175-82.
- [31] A. Magrez, S. Kasas, and N. Pasquier, *Nano Lett.* **2006**, 6, 1121-25.
- [32] H. Nomura, C. H. Tator, and M. S. Shoichet, *J. Neurotrauma* **2006**, 23, 496-507.
- [33] C. L. Cheung and C. M. Lieber, *Nature* **1998**, 394, 52-55.
- [34] A. Star, E. Tu, J. Niemann, J. P. Gabriel, C. S. Joiner, and C. Valcke, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, 103, 921-26.
- [35] B. Wieckowski, K. Kostarelos, L. Lacerda, G. Pastorin, W. E. I. Wu, J. Luangsivilay, S. Godefroy, D. Pantarotto, J. Briand, S. Muller, M. Prato, and A. Bianco, *Nat. Nanotechnol.* **2007**, 2, 108-13.
- [36] N. Wong, S. Kam, M. O. Connell, J. A. Wisdom, and H. Dai, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, 11600-05.
- [37] H. Cui, S. K. Vashist, O. K. Al-Rubeaan, and J. H. T. Luong, *Chem. Res. Toxicol.* **2010**, 23, 1131-47.
- [38] V. L. Colvin, *Nat. Biotechnol.* **2003**, 21, 1166-71.
- [39] D. Cui, F. Tian, C. S. Ozkan, M. Wang, and H. Gao, *Toxicol. Lett.* **2005**, 155, 73-85.
- [40] G. Jia, H. Wang, L. Yan, X. Wang, R. Pei, T. Yan, Y. Zhao, and X. Guo, *Environ. Sci. Technol.* **2005**, 39, 1378-83.
- [41] L. P. Zanello, B. Zhao, H. Hu, and R. C. Haddon, *Nano Lett.* **2006**, 6, 562-67.

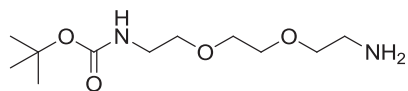
- [42] Z. Liu, X. Sun, N. Nakayama-Ratchford, and H. Dai, *ACS Nano* **2007**, 1, 50–56.
- [43] V. C. Moore, M. S. Strano, E. H. Haroz, R. H. Hauge, R. E. Smalley, J. Schmidt, and Y. Talmon, *Nano Lett.* **2003**, 3, 1379-82.
- [44] T. Raman and S. Properties, *Chem. Eur. J.* **2006**, 12, 7595-602.
- [45] P. Liu, *Eur. Polym. J.* **2005**, 41, 2693-703.
- [46] V. Georgakilas, N. Tagmatarchis, D. Pantarotto, A. Bianco, D. B. Mol, and U. P. R. CNRs, *Chem. Commun.* **2002**, 3050-51.
- [47] S. Daniel, T. Prasada, K. Sreenivasa, S. Usha, G. R. K. Naidu, H. Lee, and T. Kawai, *Sens. Actuators B* **2007**, 122, 672-82.
- [48] B. K. Gorityala, J. Ma, X. Wang, X. Liu, and X. Wang, *Chem. Soc. Rev.* **2010**, 39, 2925-34.
- [49] X. Chen, G. S. Lee, A. Zettl, and C. R. Bertozzi, *Angew. Chemie Int. Ed.* **2004**, 43, 6111-16.
- [50] X. Chen, U. C. Tam, J. L. Czapinski, G. S. Lee, D. Rabuka, A. Zettl, C. R. Bertozzi, V. Uni, M. Sciences, V. Di, and L. Berkeley, *J. Am. Chem. Soc.* **2006**, 128, 6292-93.
- [51] B. Nanotubes, P. Wu, X. Chen, N. Hu, U. C. Tam, O. Blixt, A. Zettl, and C. R. Bertozzi, *Angew. Chemie Int. Ed.* **2008**, 47, 5022-25.
- [52] I. Fern, M. Assali, M. Perni, and R. Baati, *Soft Matter* **2009**, 5, 948-50.
- [53] B. Ballesteros, G. Tobias, L. Shao, E. Pellicer, J. Nogués, E. Mendoza, and M. L. H. Green, *Small* **2008**, 4, 1501-6.
- [54] R. Meyer, J. Sloan, R. E. Dunin-Borkowski, A. I. Kirkland, M. C. Novotny, S. R. Bailey, J. L. Hutchison, and M. L. H. Green, *Science* **2000**, 289, 1324-27.

- [55] S. Y. Hong, G. Tobias, F. El Oualid, J. C. Errey, K. J. Doores, A. I. Kirkland, P. D. Nellist, M. L. H. Green, and B. G. Davis, *J. Am. Chem. Soc.* **2007**, 129, 10966-67.
- [56] M. Maggini, G. Scorrano, and M. Prato, *J. Am. Chem. Soc.* **1993**, 115, 9798-99.
- [57] V. L. Kan and J. E. Bennett, *J. Infect. Dis.* **1988**, 158, 407-14.
- [58] T. P. Hickling, H. Clark, R. Malhotra, and R. B. Sim, *J. Leukoc. Biol.* **2004**, 75, 27-33.
- [59] B. G. Davis, *Chem. Rev.* **2002**, 44, 579-601.
- [60] E. L. Mazzaferri, in *The Thyroid: A Fundamental and Clinical Text* (Eds.: by L. E. Braverman, R. D. Utiger), Lippincott-Raven, **1996**, 922-45.
- [61] L. Lacerda, A. Soundararajan, R. Singh, G. Pastorin, K. T. Al-Jamal, J. Turton, P. Frederik, M. A. Herrero, S. Li, A. Bao, D. Emfietzoglou, S. Mather, W. T. Phillips, M. Prato, A. Bianco, B. Goins, and K. Kostarelos, *Adv. Mater.* **2008**, 20, 225-30.
- [62] S. S. Onoke, T. U. Eda, K. F. Ujiwara, K. K. Uwabara, and J. Y. Ano, *Biol. Pharm. Bull.* **2011**, 34, 1338-42.
- [63] J. Kong and A. M. Cassell, *Chem. Phys. Lett.* **1998**, 567-74.
- [64] C. T. Wirth, B. C. Bayer, A. D. Gamalski, S. Esconjauregui, R. S. Weatherup, C. Ducati, C. Baehtz, J. Robertson, and S. Hofmann, *Chem. Mater.* **2010**, 24, 4633-40.
- [65] L. Shao, G. Tobias, C. G. Salzmann, A. Crossley, B. G. Davis, and M. L. H. Green, *Chem. Commun. (Camb)*. **2007**, 1, 5090-92.
- [66] A. Shvedova, V. Castranova, E. Kisin, D. Schwegler-Berry, A. Murray, V. Gandelsman, A. Maynard, P. Baron, A. Shvedova, V. Castranova, E. Kisin,

- D. Schwegler-Berry, *J. Toxixology Environ. Heal. Part A* **2003**, 66, 1909-26.
- [67] A. Nimmagadda, K. Thurston, M. U. Nollert, P. S. McFetridge *J. Biomed. Mater. Res. A* **2006**, 3, 614-25.
- [68] V. E. Kagan, Y. Y. Tyurina, V. A. Tyurin, N. V. Konduru, A. I. Potapovich, A. N. Osipov, E. R. Kisin, D. Schwegler-Berry, R. Mercer, V. Castranova, and A. A. Shvedova, *Toxicol. Lett.* **2006**, 165, 88-100.
- [69] G. Tobias, L. Shao, C. G. Salzmann, Y. Huh, and M. L. H. Green, *J. Phys. Chem. B* **2006**, 110, 22318-22.
- [70] G. Tobias, M. A. H. Ward, and M. L. H. Green, *J. Phys. Chem. C* **2009**, 113, 2653-56.
- [71] A. W. Musumeci, G. G. Silva, W. N. Martens, E. R. Wacławik, and R. L. Frost, *J. Therm. Anal. Cal.* **2007**, 88, 885-91.
- [72] M. Monthieux, *Carbon N. Y.* **2002**, 40, 1809-23.
- [73] S. Rana and J. W. Cho, *Nanoscale* **2011**, 2, 2550-56.
- [74] R. R. Slama, J., Rando, *Carbohydr. Res.* **1981**, 88, 213-21.
- [75] *J. Org. Chem. USSR* **1971**, 7, 2250-53 (English translation).
- [76] *Bull. Soc. Chim. Fran.* **1951**, 340.

5.7 Supporting information

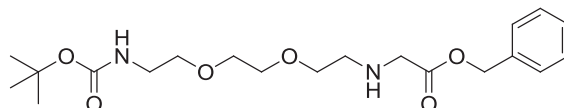
N-*tert*-butoxycarbonyl-2,2'-(ethylenedioxy)diethylamine **5.1**^[1]



To a solution of 2,2'-(ethylenedioxy)diethylamine (5.1 mL, 34 mmol, 7.4 eq.) in 1,4-dioxane (12.5 mL) at 0 °C was added a solution of di-*tert*-butyl dicarbonate (1 g, 4.6 mmol, 1 eq.) in 1,4-dioxane (12.5 mL) over 40 min at 0 °C. The reaction mixture was stirred overnight at r.t. before being concentrated under vacuum into a white slurry and recovered with H₂O (20 mL). The aqueous solution was extracted with DCM (3 x 20 mL) and the combined organic layer was backwashed with water (20 mL) and dried with MgSO₄, filtered and evaporated. The residue was purified by flash column chromatography over silica (DCM:MeOH:33% NH₃ in H₂O, 10:0.3:0.01, then MeOH:33% NH₃ in H₂O, 99:1) to afford the title compound **5.1** as a colourless viscous oil (1.08 g, 4.35 mmol, 95%). **TLC** R_f = 0.2 (DCM:MeOH:33% NH₃ in H₂O, 10:0.3:0.01); **¹H NMR** (CDCl₃, 500 MHz) δ 3.25 (br, 4 H, OCH₂CH₂O), 3.16 (br, 4 H, 2 x OCH₂CH₂N), 2.92 (br, 2 H, CH₂NHCO), 2.49 (t, *J* = 4.8 Hz, 2 H, CH₂NH₂), 1.98 (br, 2 H, CH₂NH₂), 1.07 (s, 9 H, ^{*t*}Bu); **¹³C NMR** (CDCl₃, 126 MHz) δ 156.27 (CO), 78.82 (C(CH₃)₃), 73.20 (2 x OCH₂CH₂N), 70.24 (OCH₂CH₂O), 41.63 (CH₂NH₂), 40.37 (CH₂NHCO), 28.5 (C(CH₃)₃); **MS** (ESI⁺) *m/z* = 249.33 [M+H]; **IR** (ATR) 3291.47 (NH₂), 1684.15 (CO).

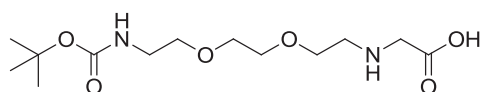
N-(N-*tert*-butoxycarbonyl-(2,2'-(ethylenedioxy)diethylamino))

benzylacetate **5.2^[2]**



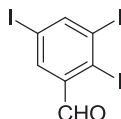
To a solution of **5.1** (500 mg, 2.0 mmol, 3 eq.) in 1,4-dioxane (1 mL) was added a solution of benzylbromoacetate (156 μ L, 0.67 mmol, 1 eq.) in 1,4-dioxane (1.5 mL) at 0 °C over 20 min. The reaction mixture was stirred overnight at r.t.. The reaction mixture was concentrated *in vacuo*, re-dissolved in EtOAc (20 mL) and washed with water (20 mL). The aqueous layer was extracted with EtOAc (2 x 20 mL). The combined organic layer was dried with MgSO₄, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography over silica (EtOAc 100%) to afford the title compound **5.2** as a pale yellow viscous oil (260 mg, 0.66 mmol, 98%). **TLC** R_f = 0.25 (EtOAc, 100%); **¹H NMR** (CDCl₃, 250 MHz) δ 7.34 (br, 5 H, CH-Ar), 5.15 (s, 2 H, CH₂Ph), 3.58 (br, 4 H, OCH₂CH₂O), 3.51 (br, 2 H, CH₂NHCO), 3.48 (br, 4 H, 2 x OCH₂CH₂N), 3.27 (dd, *J* = 10.5, 5.3 Hz, 2 H, CH₂NHCH₂), 2.81 (t, *J* = 5.2 Hz, 2 H, CH₂NHCO), 2.03 (br, 2 H, CH₂NH₂), 1.42 (s, 9 H, ^{*t*}Bu); **¹³C NMR** (CDCl₃, 63 MHz) δ 172.60 (CH₂CO), 156.27 (NHCO), 136.10 (C-Ar), 128.75, 128.99 (5 x CH-Ar), 79.51 (C(CH₃)₃), 71.12 (2 x OCH₂CH₂N), 70.66 (OCH₂CH₂O), 66.88 (CH₂Ph), 51.32 (NHCH₂CO), 49.15 (CH₂NHCH₂), 40.84 (CH₂NHCO), 28.82 (C(CH₃)₃); **MS** (ESI⁺) *m/z* = 397.48 [M+H]; **IR** (ATR) 3251.56 (NH), 2976.35 (CH), 1672.45 (CO), 1633.42 (COOR), 697.15 (arom).

N-*tert*-butoxycarbonyl-linker 1 **5.3**^[2]



To a solution of **5.2** (255 mg, 0.6 mmol, 1 eq.) in MeOH (dry, 1.5 mL) was added Pd/C (10%, 7 mg) and the mixture was stirred under hydrogen gas at r.t.. After 8 h the reaction mixture was filtered through celite, and the solvent was removed *in vacuo*. The product was triturated in diethyl ether to afford the title compound **5.3** as a colourless viscous oil (171 mg, 0.56 mmol, 87%). **TLC** R_f = 0.0 (EtOAc, 100%); **¹H NMR** (MeOD, 250 MHz) δ 3.81 (br, 2 H, CH₂NHCH₂), 3.68 (br, 6 H, NHCH₂COOH, OCH₂CH₂O), 3.57 (t, J = 5.5 Hz, 2 H, OCH₂CH₂NHCH₂), 3.28 (br, 2 H, CH₂CH₂NHCO), 3.26 (m, 2 H, CH₂CH₂NHCO), 1.46 (s, 9 H, ^{*t*}Bu); **¹³C NMR** (MeOD, 63 MHz) δ 169.85 (COOH), 157.42 (NHCO), 79.06 (C(CH₃)₃), 70.15 (CH₂CH₂NHCO), 70.06 (OCH₂CH₂O), 66.13 (CH₂CH₂NHCH₂), 48.94 (CH₂CH₂NHCH₂, NHCH₂COOH), 40.17 (CH₂CH₂NHCO), 27.85 (C(CH₃)₃); **MS** (ESI⁺) m/z = 307.36 [M+H]; **IR** (ATR) 3348.97 (OH), 2976.35 (CH), 1693.44 (CO), 1631.77 (COOH).

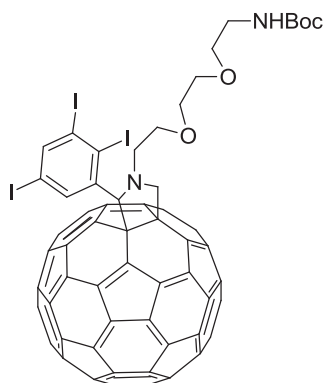
2,3,5-triiodobenzaldehyde **5.4**



2,3,5-triiodobenzyl alcohol (990 mg, 2 mmol, 1 eq.) and PCC (862 mg, 4 mmol, 2 eq.) were purged with argon gas. To the dried compounds, THF (dry, 25 mL)

was added and stirred at r.t.. After 18 h, additional PCC (431 mg, 2 mmol) was added to the reaction mixture and stirred for another 6 h under argon. The reaction mixture was filtered through celite, and the residue was concentrated *in vacuo* and purified by flash column chromatography over silica (PE:DCM, 10:1) to afford the title compound **5.4** as a white solid (513 mg, 1.06 mmol, 53%). **TLC** R_f = 0.3 (DCM:PE, 1:10); **$^1\text{H NMR}$** (CDCl_3 , 400 MHz) δ 9.87 (s, 1 H, CHO), 8.44, 8.06 (2 x d, 2 x J = 1.3 Hz, 2 H, 2 x CH-Ar); **$^{13}\text{C NMR}$** (CDCl_3 , 63 MHz) δ 196.66 (CHO), 156.66 (2 x CH-Ar), 141.47 (C-Ar), 114.43 (3 x Cl-Ar); **MS** (FI^+) m/z = 483.81 [M]; **IR** (ATR) 1678.69 (CO).

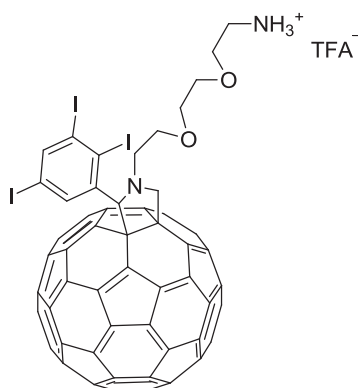
N-*tert*-butoxycarbonyl-linker 1 fullerene **5.5**



A toluene solution (200 mL) of C_{60} (400 mg, 0.56 mmol, 1 eq.), **5.3** (171 mg, 0.56 mmol, 1 eq.) and **5.4** (270 mg, 0.56 mmol, 1 eq.) was refluxed (110 °C) for 2h. After cooling to r.t., the reaction mixture was concentrated to a small volume and purified by flash column chromatography over silica (toluene:EtOAc, 9:1) to afford the title compound **5.5** as a brown solid (375 mg, 0.26 mmol, 46%). **TLC** R_f = 0.25 (toluene:EtOAc, 10:1); **$^1\text{H NMR}$** (CDCl_3 , 250 MHz) δ 8.48, 8.27 (2 x d, 2 x J = 2.0 Hz, 2 H, 2 x CH-Ar), 5.88 (s, 1 H, CH-pyrrolidine), 5.20 (d, J = 9.7 Hz, 1 H, CH_2 -

pyrrolidine), 5.06 (br, 1 H, NH), 4.46 (d, $J = 9.7$ Hz, 1 H, CH₂-pyrrolidine), 4.02 (t, $J = 5.1$ Hz, 2 H, NCH₂), 3.83 (m, 4 H, OCH₂CH₂O), 3.68 (t, $J = 5.1$ Hz, 2 H, NCH₂CH₂), 3.45-3.31 (m, 3 H, 1 x CH₂CH₂NH, CH₂NH), 2.94-3.04 (m, 1 H, CH₂CH₂NH), 1.52 (s, 9 H, ^tBu); **¹³C NMR** (CDCl₃, 63 MHz) δ 156.77-125.76 (CH-Ar, C-Ar, C-fullerene), 115.90, 113.27, 95.63 (3 x Cl-Ar), 86.58 (CH-pyrrolidine), 79.74 (C(CH₃)), 71.06, 70.88, 70.36, 69.72 (NCH₂CH₂, NCH₂CH₂, OCH₂CH₂O), 67.31 (CH₂-pyrrolidine), 52.23 (CH₂CH₂NH), 40.88 (CH₂CH₂NH), 28.97 (C(CH₃)₃); **MS** (LD⁺) m/z 1448.78 [M]; **IR** (ATR) 1697.34 (CO).

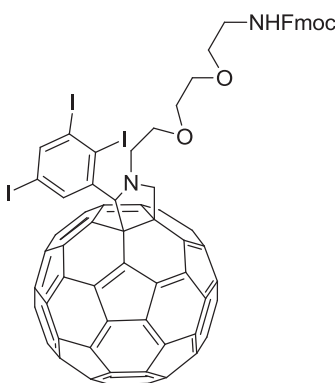
NH₂-linker 1 fullerene **5.6**



To a solution of **5.5** (375 mg, 0.26 mmol, 1 eq.) in DCM (dry, 10 mL) was added TFA (5 mL). The reaction mixture was stirred for 2.5 h at r.t.. The reaction was evaporated, sonicated in DCM and in MeOH. The solid was dried under vacuum to afford the title compound **5.6** as a brown solid (220.7 mg, 0.16 mmol, 63%). **TLC** $R_f = 0.35$ (toluene:EtOAc, 4:1); **¹H NMR** (DMSO-*d*₆, 400 MHz) δ 8.33, 8.28 (2 x br, 2 H, 2 x CH-Ar), 7.79 (br, 2 H, NH₂), 5.87 (br, 1 H, CH-pyrrolidine), 5.26 (d, $J = 9.2$ Hz, 1 H, CH₂-pyrrolidine), 4.51 (d, $J = 9.1$ Hz, 1 H, CH₂-pyrrolidine), 3.89 (m, 2 H, NCH₂CH₂), 3.74 (m, 4 H, OCH₂CH₂O), 3.68 (m, 2 H, NCH₂CH₂),

2.99-2.88 (m, 4 H, OCH₂CH₂NH); ¹³C NMR (DMSO-d₆, 101 MHz) δ 156.94-134.89 (CH-Ar, C-Ar, C-fullerene), 116.36, 114.35, 95.93 (3 x Cl-Ar), 84.98 (CH-pyrrolidine), 69.80, 69.12, 68.66 (NCH₂CH₂, NCH₂CH₂, OCH₂CH₂O), 65.53 (CH₂-pyrrolidine), 50.76 (CH₂CH₂NH₂), 35.76 (CH₂CH₂NH₂); MS (LD⁺) m/z 1348.67 [M]; IR (ATR) 3271.00 (NH₂).

N-fluorenylmethyloxycarbonyl-linker 1 fullerene **5.7**



A solution of Fmoc chloride (3 mg, 11.6 μmol, 1.2 eq.) in DCM (dry, 2 mL) was added dropwise to a suspension of **5.6** (14 mg, 9.57 μmol, 1 eq.) in DCM (dry, 5 mL) at 0 °C. To the mixture was added DIPEA (3.5 μL, 20 μmol, 2 eq.) and it was stirred for 1.5 h at 0 °C. The reaction was stirred overnight at r.t., then diluted with DCM (10 mL), washed with water (2 x 10 mL), dried over MgSO₄ and concentrated *in vacuo*. The residue was sonicated in methanol and recovered by centrifugation, to afford the title compound **5.7** as a brown solid (12.7 mg, 8.08 μmol, 84%). TLC R_f = 0.7 (toluene:EtOAc, 4:1); ¹H NMR (CDCl₃, 400 MHz) δ 8.44, 8.23 (2 x br, 2 H, 2 x CH-Ar), 7.78 (t, *J* = 7.4 Hz, 2 H, H-4-Fmoc, H-5-Fmoc), 7.62 (t, *J* = 7.1 Hz, 2 H, H-1-Fmoc, H-8-Fmoc), 7.41 (t, *J* = 7.3 Hz, 2 H, H-3-Fmoc, H-6-Fmoc), 7.33 (t, *J* = 7.2 Hz, 2 H, H-2-Fmoc, H-7-Fmoc), 5.82 (br, 1 H, CH-

pyrrolidine), 5.31 (br, 1 H, NH), 5.17 (d, $J = 9.6$ Hz, 1 H, CH₂-pyrrolidine), 4.41 (br, 3 H, 1 x CH₂-pyrrolidine, CH₂Fmoc), 4.24 (t, $J = 6.5$ Hz, 1 H, CHCH₂-Fmoc), 3.96 (m, 2 H, NCH₂), 3.79 (m, 4 H, OCH₂CH₂O), 3.67 (m, 2 H, NCH₂CH₂), 3.46, 3.30 (2 x m, 4 H, CH₂CH₂NH); ¹³C NMR (CDCl₃, 101 MHz) δ 159.99-156.42 (CH-Ar, C-Ar, C-fullerene, CO), 115.29, 112.47, 94.91 (3 x Cl-Ar), 86.26 (CH-pyrrolidine), 70.62, 70.20, 69.70, 69.14 (NCH₂CH₂, OCH₂CH₂O, CH₂Fmoc), 66.62 (CH₂-pyrrolidine), 51.88 (CH₂CH₂NH), 47.27 (CHCH₂-Fmoc), 40.99 (CH₂CH₂NH); MS (LD⁺) m/z 1570.91 [M]; IR (ATR) 1665.56 (CO).

Fmoc numbering on N-fluorenylmethyloxycarbonyl-linker 1 fullerene

N-fluorenylmethyloxycarbonyl-linker 1 fullerene **5.7** (12.7 mg, 8.08 μ mol) was treated with DMSO:DMF:DBU (25:24:1 solution, 2 mL) at r.t. to cleave the Fmoc group. After 2 h, TLC (EtOAc:MeOH, 8:2) indicated the formation of the product ($R_f = 0.0$) with complete consumption of the starting material ($R_f = 0.5$). 20 μ L of the crude were added to 1000 μ L of MeOH. The sample was centrifuged and the absorbance of the solution measured by UV at 295 nm ($A = 0.61425$).

The functionalisation of the fullerene obtained was 0.482 mmol/g.

Purification of SWNTs

Triturated SWNTs were placed in a furnace under argon and heated up to 900 °C while steam was fluxed through the chamber. The temperature was raised by 300 °C/min till 900 °C, where the SWNTs remained for 12 h. The system was

slowly cooled at a rate of 60 °C/h down to 500°C and at 300 °C/h down to 25 °C to allow the tips of the SWNTs to close^[1,2] (Fig. S1).

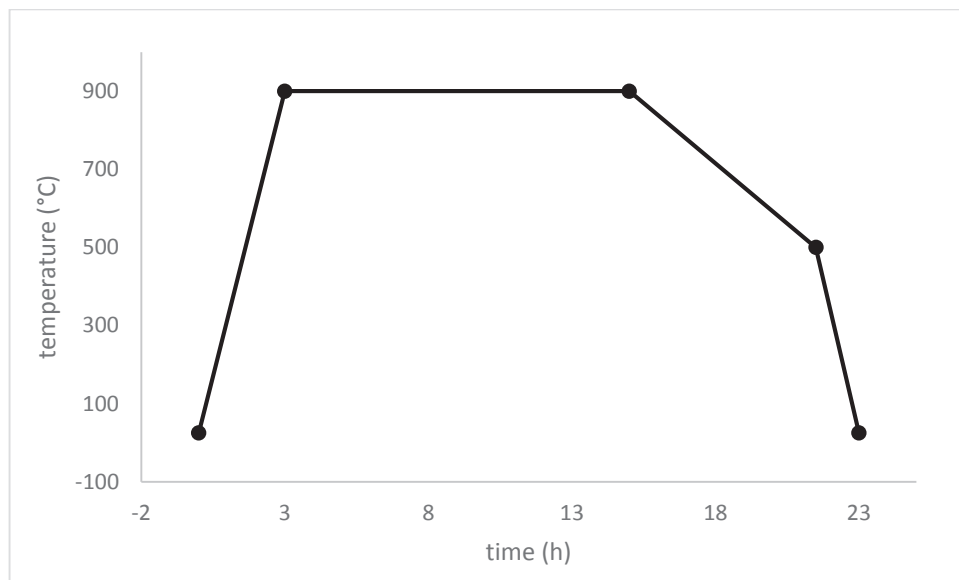
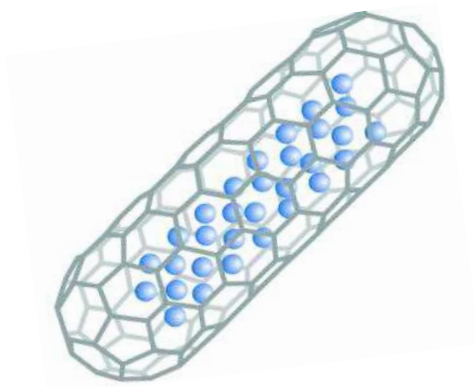


Figure S1. Thermal gradient for the purification of SWNTs.

SWNTs were recovered with HCl 6 M and refluxed at 110 °C for 6 h to eliminate any residual Iron. SWNTs were filtered and washed with H₂O till neutral pH, and finally dried at 60 °C for 16 h.

Nal@SWNTs 5.8



Steam-purified SWNTs (1 mg) were loaded in a quartz ampoule with NaI solution (10^{-5} M, 100 μ L), dried and sealed under high vacuum with an oxygen-propane flame. The ampoule was heated in a furnace up to 900 °C using the following temperature ramp (Fig. S2):

1. from 0 °C to 900 °C in 3 h (300 °C/h)
2. 900 °C for 4 h
3. from 900 °C to 500 °C in 6.5 h (60 °C/h)
4. from 500 °C to r.t. in 2 h (300 °C/h)

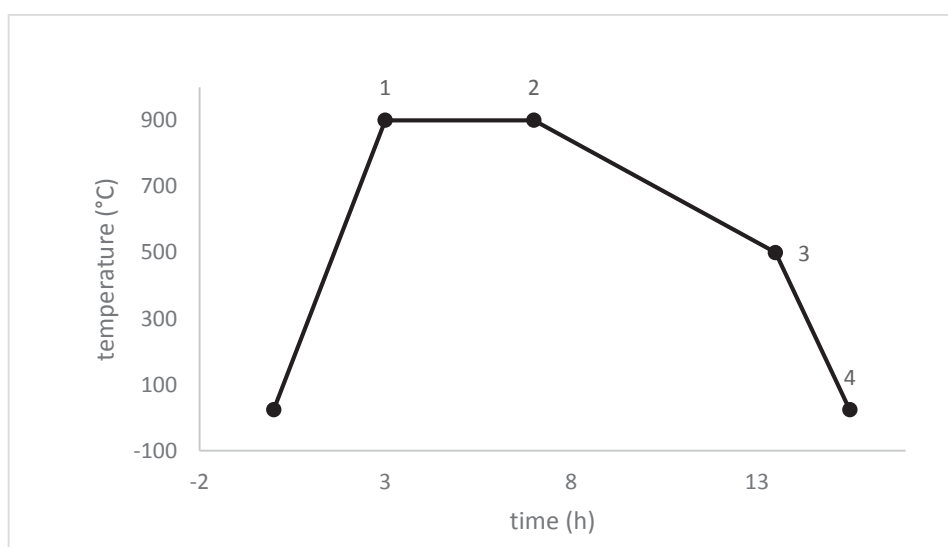
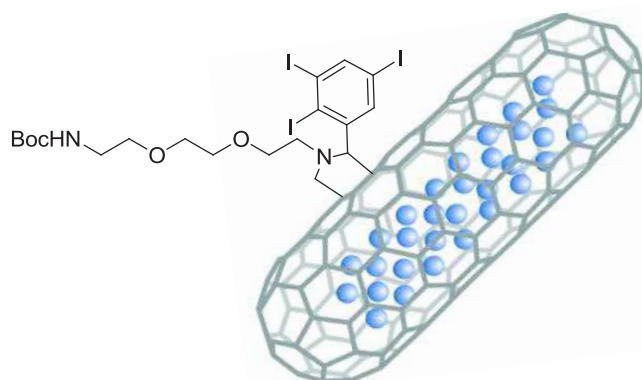


Figure S2. Thermal gradient for the filling of SWNTs with NaI.

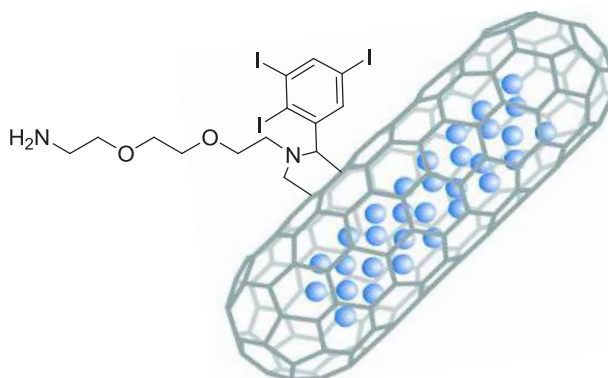
The sample was recovered from the ampoule with H₂O (1 mL), sonicated and filtered.

N-*tert*-butoxycarbonyl-linker 1 NaI@SWNTs **5.9**



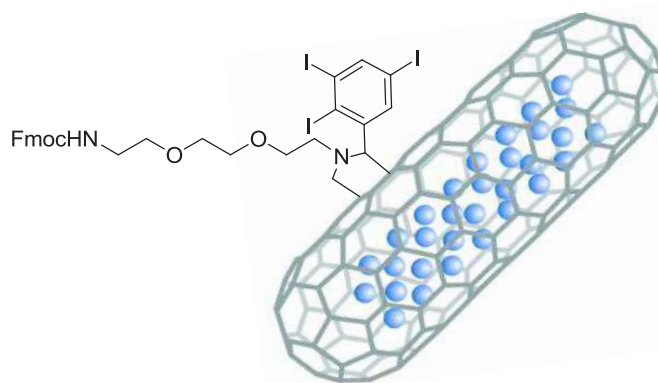
A suspension of **5.8** (8 mg) in DMF (dry, 3 mL) was sonicated for 2 min, before a solution of **5.3** (6.13 mg, 0.02 mmol, 1 eq.) and **5.4** (9.6 mg, 0.02 mmol, 1 eq.) in DMF (dry, 1 mL) was added. The reaction was refluxed at 130 °C for 4 days, then cooled to r.t. and filtered. The solid was washed with DMF, MeOH and dried under vacuum to afford the title compound **5.9** as a black solid (8.8 mg).

NH₂-linker 1 NaI@SWNTs **5.10**



A suspension of **5.9** (12.1 mg) in DCM (dry, 5 mL) was sonicated for 2 min, before TFA (2.5 mL) was added. The reaction was stirred at r.t. for 24 h, then evaporated, recovered with MeOH, filtered and washed MeOH. The solid obtained was dried under vacuum to afford the title compound **5.10** as a black solid (8.8 mg). **EA** (C,H,N) 65.06 ,1.00, 2.61.

N-fluorenylmethyloxycarbonyl-linker 1 Ial@SWNTs **5.11**



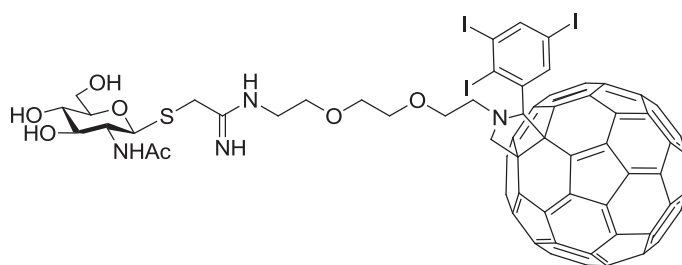
A solution of Fmoc chloride (1 mg, 3.86 μmol , 1 eq.) in DCM (0.5 mL) was added dropwise to a suspension of **5.10** (1 mg) in DCM (0.5 mL) at 0 °C. DIPEA (1.5 μL , 8.61 μmol , 2.2 eq.) was added and the reaction was stirred at r.t. for 16 h. The reaction mixture was centrifuged, the supernatant discarded and the solid recovered with MeOH, filtered and washed with MeOH. The solid obtained was dried under vacuum to afford the title compound **5.11** as a black solid (0.95 mg). **EA** (C,H,N) 66.32 ,0.90, 2.33.

Fmoc numbering on N-fluorenylmethyloxycarbonyl-linker 1 NaI@SWNTs

5.11 (0.95 mg) were treated with DMSO:DMF:DBU (25:24:1 solution, 1 mL) at r.t. to cleave the Fmoc group. After 2 h, 20 μ L of the crude were added to 1000 μ L of MeOH. The sample was centrifuged and the absorbance of the solution measured by UV at 295 nm ($A = 0.031285$).

The functionalisation of the nanotubes obtained was 0.162 mmol/g (literature value = 0.17 mmol/g).

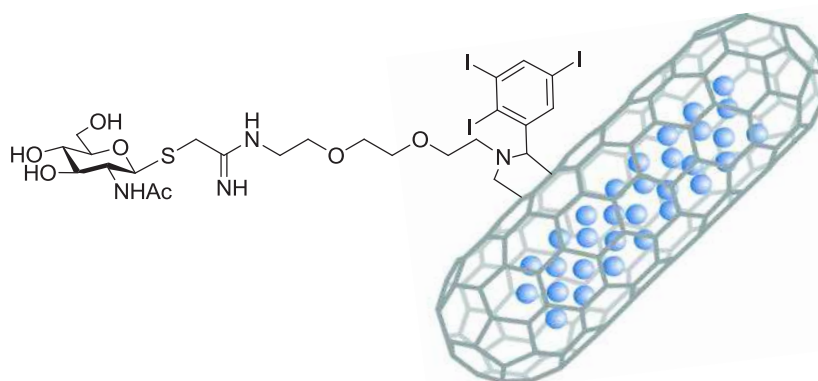
GlcNAc-linker 1 fullerene **5.12**



To a solution of **5.6** (14 mg, 0.01 mmol, 1 eq.) in MeOH:DMSO (2:1, 1.5 mL) was added **2.80** (6 mg, 0.02 mmol, 2 eq.). The reaction was stirred at r.t. for 30 min, then filtered, washed with MeOH and Et₂O, to afford the title compound **5.12** as a black solid (10.4 mg, 6.4 μ mol, 64%). **TLC** $R_f = 0.2$ (toluene:EtOAc, 4:1); **¹H NMR** (MeOD, 400 MHz) δ 8.30, 8.27 (2 x br, 2 H, 2 x CH-Ar), 5.85 (br, 1 H, CH-pyrrolidine), 5.14 (d, $J = 9.2$ Hz, 1 H, CH₂-pyrrolidine), 4.64 (br, 1 H, H-1), 4.48 (d, $J = 9.1$ Hz, 1 H, CH₂-pyrrolidine), 3.99-3.80 (m, 2 H, H-2, H-4), 3.72 (m, 2 H, NCH₂CH₂), 3.69-3.62 (m, 10 H, OCH₂CH₂O, SCH₂, H-3, H-5, H-6), 3.58 (m, 2 H, NCH₂CH₂), 2.97-2.79 (m, 4 H, OCH₂CH₂NH), 1.93 (Ac); **¹³C NMR** (MeOD, 101 MHz) δ 166.44-133.21 (CH-Ar, C-Ar, C-fullerene, CO, CNH), 116.36, 114.35,

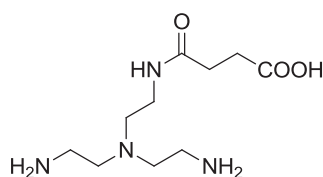
95.93 (3 x Cl-Ar), 84.98 (CH-pyrrolidine), 82.76 (C-1), 78.94 (C-5), 72.59 (C-3), 69.80, 69.12, 68.66 (NCH₂CH₂, NCH₂CH₂, OCH₂CH₂O), 68.37 (C-2), 65.53 (CH₂-pyrrolidine), 61.11 (C-6), 50.99 (C-4), 50.76 (CH₂CH₂NH), 35.76 (CH₂CH₂NH), 22.09 (Ac); **MS** (LD⁺) m/z 1624.98 [M]; **IR** (ATR) 1728.66 (COOR), 1656.27 (CONH).

GlcNAc-linker 1 NaI@SWNTs **5.13**



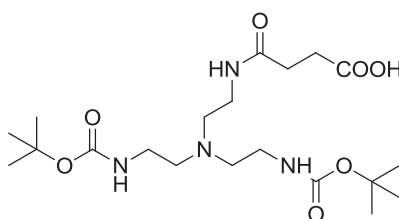
A solution of **2.80** (6 mg, 0.02 mmol, 1 eq.) in MeOH:DMSO (2:1, 1.5 mL) was added to a suspension of **5.10** (1 mg) in DCM (0.5 mL). The reaction was stirred at r.t. for 30 min. The reaction mixture was centrifuged, the supernatant discarded and the solid was recovered with MeOH, filtered and washed with MeOH. The solid obtained was dried under vacuum to afford the title compound **5.13** as a black solid (0.95 mg). **EA** (C,H,N) 67.01 ,1.01, 3.04.

4-(2-(bis(2-aminoethyl)amino)ethylamino)oxobutanoic acid 5.14



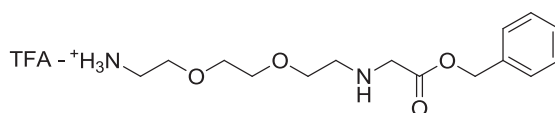
To a solution of tris(2-aminoethyl)amine (6 mL, 41 mmol, 5 eq.) in ACN (200 mL) at 0 °C was added, dropwise, a solution of succinic anhydride (822 mg, 8.2 mmol, 1 eq.) in ACN (100 mL). The reaction was stirred at r.t. for 2 h, before the solution was poured away and the precipitate recovered with MeOH, and finally dried under vacuum to afford the title compound **5.14** as a yellow oil (2.0 g, 8 mmol, 98%). **TLC** R_f = 0.1 (WIPE, 1:2:2); **¹H NMR** (MeOD, 400 MHz) δ 3.30 (t, J = 4.6 Hz, 2 H, CH₂NH), 2.86 (br, 4 H, 2 x CH₂NH₂), 2.64-2.59 (m, 6 H, 3 x CH₂N), 2.45 (br, 4 H, COCH₂CH₂CO); **¹³C NMR** (MeOD, 101 MHz) δ 175.10, 179.93 (2 x CO), 54.25, 55.12 (3 x CH₂N), 38.28, 38.53 (CH₂NH, 2 x CH₂NH₂), 32.75, 33.35 (COCH₂CH₂CO); **MS** (ESI⁺) m/z = 247.31 [M+H]; **HRMS** (ESI⁺) found 247.1763 [M+H], C₁₀H₂₃N₄O₃ requires 247.1765; **IR** (ATR) 3271.00 (OH), 1703.87 (COOH), 1697.90 (CONH).

4-(2-(bis(2-*tert*-butoxycarbonyl-aminoethyl)amino)ethylamino)oxobutanoic acid 5.15



To a solution of **5.14** (343 mg, 1.39 mmol, 1 eq.) in dioxane (15 mL) was added di-*tert*-butyl dicarbonate (611 mg, 2.8 mmol, 2 eq.). The reaction mixture was stirred for 4 h at r.t.. The reaction was evaporated, and the residue purified by flash column chromatography over silica (DCM:MeOH, 10:0.2 to 10:2), to afford the title compound **5.15** as a white solid (246.8 mg, 0.55 mmol, 40%). **TLC** R_f = 0.85 (DCM:MeOH, 10:1); **^1H NMR** (CDCl_3 , 250 MHz) δ 3.27 (br, 2 H, $\text{CH}_2\text{NHCOCH}_2$), 3.13 (br, 4 H, 2 x CH_2NHCOO), 2.52 (m, 10 H, 3 x CH_2N , $\text{COCH}_2\text{CH}_2\text{CO}$), 1.41 (s, 18 H, 2 x $t\text{Bu}$); **^{13}C NMR** (CDCl_3 , 126 MHz) δ 172.69, 157.98, (2 x COCH_2), 156.63 (2 x COOC), 81.28 ($\text{C}(\text{CH}_3)_3$), 79.46, 79.22 (3 x CH_2N), 39.10, 36.86 (3 x CH_2NH), 31.66, 31.02 ($\text{COCH}_2\text{CH}_2\text{CO}$), 28.40 ($\text{C}(\text{CH}_3)_3$); **MS** (ESI^+) m/z = 447.55 [$\text{M}+\text{H}$]; **HRMS** (ESI^+) found 447.2819 [$\text{M}+\text{H}$], $\text{C}_{20}\text{H}_{39}\text{N}_4\text{O}_7$ requires 447.2813; **IR** (ATR) 3277.19 (OH), 1708.15 (COOH), 1699.62 (CONH), 1692.96 (COOC).

N-(2,2'-(ethylenedioxy)diethylamino) benzylacetate **5.16**

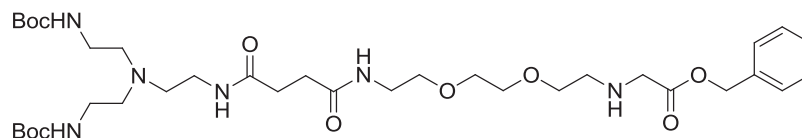


To a solution of **5.2** (1.1056 g, 2.79 mmol, 1 eq.) in DCM (5.6 mL) was added TFA (2.4 mL). The reaction was stirred at r.t. for 1 h, before being evaporated under vacuum to afford the title compound **5.16** as a yellow oil (1.08 g, 2.79 mmol, quant.). **TLC** R_f = 0.0 (EtOAc, 100%); **^1H NMR** (CDCl_3 , 500 MHz) δ 7.35-7.24 (br, 5 H, CH-Ar), 5.15 (br, 2 H, CH_2Ph), 3.93 (br, 2 H, NHCH_2CO), 3.72 (br, 2 H, $\text{OCH}_2\text{CH}_2\text{NH}_2$), 3.66 (br, 2 H, $\text{CH}_2\text{CH}_2\text{NH}$), 3.59 (br, 4 H, $\text{OCH}_2\text{CH}_2\text{O}$), 3.31 (br,

2 H, OCH₂CH₂NH), 3.22 (br, 2 H, OCH₂CH₂NH₂); **¹³C NMR** (CDCl₃, 126 MHz) δ 166.16 (CO), 133.83 (C-Ar), 118.77, 116.48, 114.19, 111.91 (5 x CH-Ar), 69.73, 69.85 (OCH₂CH₂O), 68.78 (COOCH₂Ar), 64.99, 65.73 (2 x OCH₂CH₂N), 47.89 (NHCH₂CO), 47.76 (CH₂CH₂NH), 40.10 (CH₂NH₂); **MS** (ESI⁺) m/z = 297.37 [M+H]; **HRMS** (ESI⁺) found 297.1811 [M+H], C₁₅H₂₅N₂O₄ requires 297.1809; **IR** (ATR) 3274.78 (NH₂), 1619.62 (CO), 720.61 (arom).

N-(4-(2-(bis(2-*tert*-butoxycarbonyl-aminoethyl)amino)ethylamino)oxobutanoyl)

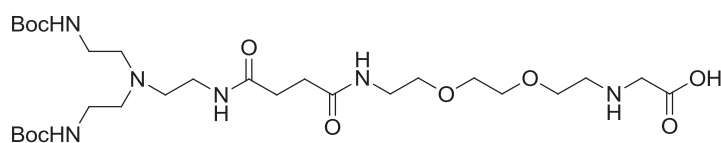
N-(2,2'-(ethylenedioxy)diethylamino) benzylacetate **5.17**



To a solution of **5.15** (1.230 g, 2.75 mmol, 1 eq.) in DMF (55 mL), were added HATU (2.091 g, 5.5 mmol, 2 eq.), HOBt (743 mg, 5.5 mmol, 2 eq.) and DIPEA (960 μL, 5.5 mmol, 2 eq.). Finally, **5.16** (978.8 mg, 3.3 mmol, 1.2 eq.) was added and the mixture stirred at r.t. for 6 h. The reaction was evaporated to a small volume, recovered with DCM (150 mL) and washed with H₂O (200 mL) and Brine (200 mL). The organic layer was dried over MgSO₄ and evaporated under vacuum, to afford the title compound **5.17** as a yellow oil (1.8455 g, 2.55 mmol, 92%). **TLC** R_f = 0.55 (DCM:MeOH, 9:1); **¹H NMR** (DMSO-d₆, 400 MHz) δ 7.86 (s, 1 H, NH), 7.46-7.31 (m, 5 H, CH-Ar), 6.79 (s, 1 H, NH), 5.23 (s, 2 H, CH₂Ph), 4.02 (s, 2 H, COCH₂NH), 3.73-3.64 (m, 2 H, OCH₂CH₂NHCO), 3.53 (dt, J = 7.9, 4.8

Hz, 6 H, OCH₂CH₂O, OCH₂CH₂NHCH₂), 3.38 (t, *J* = 6.0 Hz, 2 H, OCH₂CH₂NHCH₂), 3.25-3.07 (m, 6 H, COCH₂CH₂CO, NCH₂CH₂NHCOCH₂), 3.02 (m, 2 H, NCH₂CH₂NHCOCH₂), 2.60 (m, 4 H, NCH₂CH₂NHCOO), 2.31 (br, 4 H, 2 x CH₂NHCOO), 1.37 (s, 18 H, 2 x ^tBu); ¹³C NMR (DMSO-d₆, 101 MHz) δ 171.98 (COCH₂CH₂CO), 167.62 (CH₂COO), 156.15 (2 x NHCOO), 135.69 (C-Ar), 128.95, 128.85, 128.72 (5 x CH-Ar), 78.22 (C(CH₃)₃), 70.03 (OCH₂CH₂O), 69.82 (OCH₂CH₂NHCH₂), 69.53 (OCH₂CH₂NHCH₂), 67.25 (CH₂Ph), 66.44 (OCH₂CH₂NHCOCH₂), 53.86 (NCH₂CH₂NHCOO), 47.69 (NHCH₂CO), 46.87 (NCH₂CH₂NHCOCH₂), 38.94 (COCH₂CH₂CO), 37.94 (NCH₂CH₂NHCOCH₂), 31.06 (NCH₂CH₂NHCOO), 28.67 (C(CH₃)₃); MS (ESI⁺) *m/z* = 725.89 [M+H]; HRMS (ESI⁺) found 725.4439 [M+H], C₃₅H₆₁N₆O₁₀ requires 725.4444; IR (ATR) 3307.05 (NH), 2975.23 (CH), 1693.88 (CONH), 1650.52 (COCH₂Ph), 719.16 (arom).

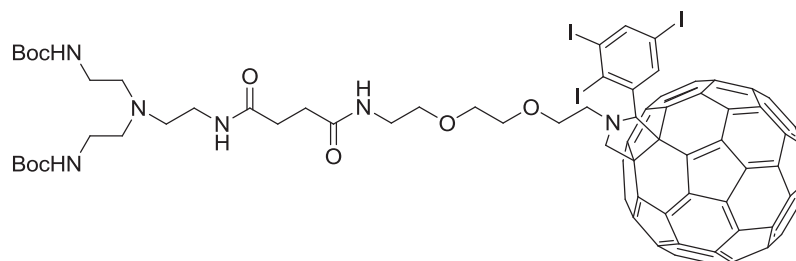
N-di(*tert*-butoxycarbonyl)-linker **5.18**



To Pd/C (10 mg, 10 wt. % loading) under hydrogen, was added a solution of **5.17** (386 mg, 0.53 mmol, 1 eq.) in MeOH (dry, 3 mL). The reaction mixture was stirred at r.t. for 5 h, filtered over celite and evaporated under vacuum. The crude was purified by flash column chromatography on silica (DCM:MeOH, 9:1 to 0:1) to afford the title compound **5.18** as a colourless oil (220.4 mg, 0.35 mmol, 66%).

TLC R_f = 0.0 (DCM:MeOH, 8:2); **$^1\text{H NMR}$** (DMSO- d_6 , 400 MHz) δ 7.98 (d, J = 5.8 Hz, 1 H, NH), 6.85 (s, 1 H, NH), 3.80-3.62 (m, 4 H, $\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$, $\text{OCH}_2\text{CH}_2\text{NHCH}_2$), 3.60-3.46 (m, 6 H, $\text{OCH}_2\text{CH}_2\text{O}$, $\text{OCH}_2\text{CH}_2\text{NHCH}_2$), 3.39 (t, J = 5.8 Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$), 3.19 (m, 4 H, $\text{COCH}_2\text{CH}_2\text{CO}$), 3.15-3.09 (m, 2 H, NHCH_2COOH), 3.06 (m, 4 H, $\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 2.72 (m, 4 H, 2 x $\text{NCH}_2\text{CH}_2\text{NHCOO}$), 2.31 (m, 4 H, 2 x $\text{NCH}_2\text{CH}_2\text{NHCOO}$), 1.37 (s, 18 H, 2 x ^tBu); **$^{13}\text{C NMR}$** (DMSO- d_6 , 101 MHz) δ 180.63 (COOH), 172.15, 171.99 (2 x COCH_2), 156.15 (2 x NHCOO), 78.33 ($\text{C}(\text{CH}_3)_3$), 72.82 ($\text{OCH}_2\text{CH}_2\text{O}$), 69.83 ($\text{OCH}_2\text{CH}_2\text{NHCH}_2$), 69.56 ($\text{OCH}_2\text{CH}_2\text{NHCH}_2$), 66.17 ($\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$), 53.34 ($\text{NCH}_2\text{CH}_2\text{NHCOO}$), 47.83 (NHCH_2COOH), 46.44 ($\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 38.92 ($\text{COCH}_2\text{CH}_2\text{CO}$), 36.22 ($\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 31.08 ($\text{NCH}_2\text{CH}_2\text{NHCOO}$), 28.67 ($\text{C}(\text{CH}_3)_3$); **MS** (ESI^-) m/z = 633.76 $[\text{M}-\text{H}]$; **HRMS** (ESI^-) found 633.3833 $[\text{M}-\text{H}]$, $\text{C}_{28}\text{H}_{53}\text{N}_6\text{O}_{10}$ requires 633.3829; **IR** (ATR) 3342.83 (NH), 1634.52 (CONH), 1700.11 (COOH).

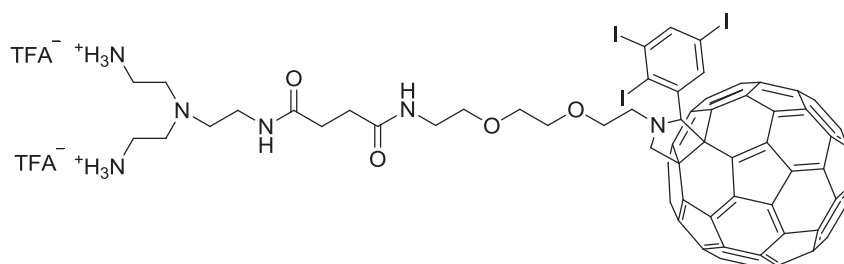
N-di(*tert*-butoxycarbonyl)-linker 2 fullerene **5.19**



To a solution of **5.18** (220.4 mg, 0.35 mmol, 1 eq.) in toluene (10 mL), were added C_{60} (250 mg, 0.35 mmol, 1 eq.) and **5.4** (168 mg, 0.35 mmol, 1 eq.). The mixture was sonicated for 15 min and refluxed at 110 °C for 20 h. The reaction was cooled

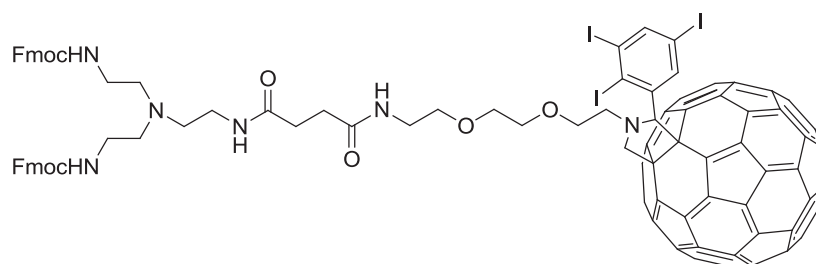
to r.t. and filtered. The solid was washed with toluene to afford the title compound **5.19** as a brown solid (337.7 mg, 0.26 mmol, 54%). **TLC** R_f = 0.25 (EtOAc:tol, 1:3); **$^1\text{H NMR}$** (CDCl_3 , 250 MHz) δ 8.49 (d, J = 2.1 Hz, 1 H, CH-Ar), 8.28 (d, J = 2.1 Hz, 1 H, CH-Ar), 7.56 (s, 1 H, NH), 6.69 (s, 1 H, NH), 5.89 (s, 1 H, CH-pyrrolidine), 5.71 (s, 1 H, NH), 5.21 (d, J = 9.6 Hz, 1 H, CH_2 -pyrrolidine), 5.10 (s, 1 H, NH), 4.46 (d, J = 9.7 Hz, 1 H, CH_2 -pyrrolidine), 4.02 (t, J = 5.4 Hz, 2 H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.85-3.80 (m, 6 H, $\text{OCH}_2\text{CH}_2\text{O}$, $\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$), 3.67 (t, J = 5.2 Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{N}$), 3.53 (t, J = 5.2 Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$), 3.32 (m, 4 H, $\text{COCH}_2\text{CH}_2\text{CO}$), 3.17 (br, 4 H, $\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 2.59 (br, 8 H, 2 x $\text{NCH}_2\text{CH}_2\text{NHCOO}$), 1.51 (s, 18 H, 2 x ^tBu); **$^{13}\text{C NMR}$** (CDCl_3 , 126 MHz) δ 177.60-118.02 (2 x COO, C-fullerene, C-Ar, CH-Ar, 2 x COCH_2), 112.65, 110.39, 94.93 (3 x Cl-Ar), 86.39 (CH-pyrrol), 79.42 ($\text{C}(\text{CH}_3)_3$), 70.68 ($\text{OCH}_2\text{CH}_2\text{O}$), 70.27 ($\text{OCH}_2\text{CH}_2\text{N}$), 69.91 ($\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$), 69.28 ($\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$), 66.95 (CH_2 -pyrrolidine), 56.22 ($\text{NCH}_2\text{CH}_2\text{NHCOO}$), 53.77, 51.85 ($\text{COCH}_2\text{CH}_2\text{CO}$), 52.34 ($\text{NCH}_2\text{CH}_2\text{NHCOO}$), 39.62 ($\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 33.70 ($\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 28.50 ($\text{C}(\text{CH}_3)_3$); **MS** (LD^+) m/z 1777.19 [M]; **IR** (ATR) 1677.92 (CO).

(NH_2)₂-linker 2 fullerene **5.20**



To a solution of **5.19** (20 mg, 0.011 mmol, 1 eq.) in DCM (2 mL), was added TFA (1 mL). The reaction was stirred at r.t. for 2 h, then evaporated under vacuum, recovered with DMSO (1 mL), precipitated with ACN (3 mL) and washed with ACN (10 mL) and acetone (10 mL) to afford the title compound **5.20** as a brown solid (17 mg, 0.011 mmol, quant). **TLC** R_f = 0.15 (DCM:MeOH, 10:1); **^1H NMR** (CDCl_3 , 400 MHz) δ 8.33 (d, J = 1.9 Hz, 1 H, CH-Ar), 8.28 (d, J = 1.9 Hz, 1 H, CH-Ar), 7.96 (br, 1 H, NH), 7.69 (br, 1 H, NH), 7.61 (br, 4 H, 2 x NH_2), 5.88 (s, 1 H, CH-pyrrolidine), 5.26 (d, J = 9.6 Hz, 1 H, CH_2 -pyrrolidine), 4.51 (d, J = 9.7 Hz, 1 H, CH_2 -pyrrolidine), 3.92-3.85 (m, 2 H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.74-3.68 (m, 4 H, $\text{OCH}_2\text{CH}_2\text{O}$), 3.49 (t, J = 5.8 Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$), 3.25-2.87 (m, 6 H, $\text{OCH}_2\text{CH}_2\text{N}$, $\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$, $\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 2.87 (br, 4 H, 2 x $\text{NCH}_2\text{CH}_2\text{NH}_2$), 2.61 (br, 4 H, 2 x $\text{NCH}_2\text{CH}_2\text{NH}_2$), 2.45 (m, 2 H, $\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 2.31 (m, 4 H, $\text{COCH}_2\text{CH}_2\text{CO}$); **^{13}C NMR** (CDCl_3 , 126 MHz) δ 171.69-139.00 (C-fullerene, C-Ar, CH-Ar, 2 x COCH_2), 116.36, 114.37, 95.96 (3 x Cl-Ar), 84.96 (CH-pyrrolidine), 69.86, 69.72 ($\text{OCH}_2\text{CH}_2\text{O}$), 69.27 ($\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$), 69.14 ($\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$), 68.67 ($\text{OCH}_2\text{CH}_2\text{N}$), 65.54 (CH_2 -pyrrolidine), 52.76 ($\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 50.94 (2 x $\text{NCH}_2\text{CH}_2\text{NH}_2$), 38.69 ($\text{OCH}_2\text{CH}_2\text{N}$), 36.77 (2 x $\text{NCH}_2\text{CH}_2\text{NH}_2$), 36.22 ($\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 30.54, 30.28 ($\text{COCH}_2\text{CH}_2\text{CO}$); **MS** (LD^+) m/z 1576.96 [M]; **IR** (ATR) 3281.42 (NH_2), 1624.85 (CO).

N-di(fluorenylmethyloxycarbonyl)-linker 2 fullerene **5.21**



To a solution of **5.20** (8 mg, 5 μ mol, 1 eq) in DMF (dry, 1.5 mL) at 0 °C, was added dropwise a solution of Fmoc chloride (3.15 mg, 12.2 μ mol, 2.4 eq.) in DCM (dry, 1 mL). To the mixture was added DIPEA (3.7 μ L, 21 μ mol, 4.2 eq) and stirred for 2 h at 0 °C, then at r.t. for 16 h. The mixture was precipitated and washed with MeOH to afford the title compound **5.21** as a brown solid (8.6 mg, 4.25 μ mol, 85%). **TLC** R_f = 0.25 (DCM:MeOH, 10:1); **^1H NMR** (DMSO- d_6 , 500 MHz) δ 8.32 (m, 1 H, CH-Ar), 8.17 (m, 1 H, CH-Ar), 7.86 (d, J = 7.5 Hz, 4 H, 4 x CH-Ar-Fmoc), 7.65 (d, J = 7.4 Hz, 4 H, 4 x CH-Ar-Fmoc), 7.39 (t, J = 7.2 Hz, 4 H, 4 x CH-Ar-Fmoc), 7.34-7.25 (m, 4 H, 4 x CH-Ar-Fmoc), 6.53 (s, 1 H, NH), 5.85 (s, 1 H, CH-pyrrolidine), 5.76 (s, 1 H, NH), 5.21 (d, J = 9.8 Hz, 1 H, CH₂-pyrrolidine), 4.45 (d, J = 9.6 Hz, 1 H, CH₂-pyrrolidine), 4.27 (m, 4 H, 2 x CH₂-Fmoc), 4.18 (m, 2 H, 2 x CH-Ar-Fmoc), 3.91-3.81 (m, 2 H, OCH₂CH₂N), 3.66 (br, 4 H, OCH₂CH₂O), 3.47 (t, J = 5.9 Hz, 2 H, OCH₂CH₂N), 3.20 (d, J = 6.0 Hz, 2 H, OCH₂CH₂NH), 3.02 (m, 6 H, OCH₂CH₂NH, COCH₂CH₂CO), 2.49 (dd, J = 9.9, 8.2 Hz, 4 H, NCH₂CH₂NHCOCH₂), 2.30 (m, 8 H, 2 x NCH₂CH₂NHCOO); **^{13}C NMR** (DMSO- d_6 , 126 MHz) δ 177.23-136.22 (CH-Ar, C-Ar, C-fullerene, 4 x CO), 115.33, 113.45, 96.81 (3 x Cl-Ar), 85.34 (CH-pyrrolidine), 70.22, 70.11 (OCH₂CH₂O), , 69.70, 69.14 (2 x CH₂Fmoc), 69.07 (OCH₂CH₂NHCOCH₂), 68.14

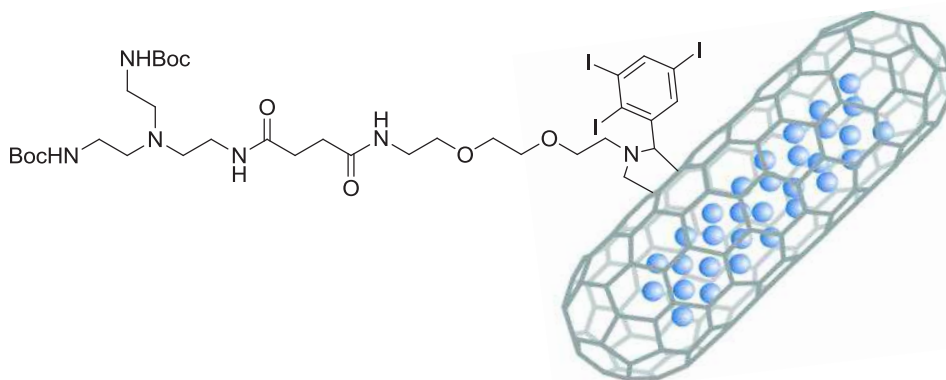
(OCH₂CH₂NHCOCH₂), 67.67 (OCH₂CH₂N), 66.42 (CH₂-pyrrolidine), 53.77 (NCH₂CH₂NHCOCH₂), 51.64 (2 x NCH₂CH₂NHCOO), 38.56 (OCH₂CH₂N), 36.49 (2 x NCH₂CH₂NHCOO), 35.98 (NCH₂CH₂NHCOCH₂), 30.51, 30.12 (COCH₂CH₂CO); **MS** (LD⁺) m/z 2021.43 [M]; **IR** (ATR) 3266.78 (NH₂), 1633.44 (CO).

Fmoc numbering on N-di(fluorenylmethyloxycarbonyl)-linker 2 fullerene

N-di(fluorenylmethyloxycarbonyl)-linker 2 fullerene **5.21** (8.6 mg, 4.25 μmol) was treated with DMSO:DMF:DBU (25:24:1 solution, 2 mL) at r.t. to cleave the Fmoc group. After 2 h, TLC (EtOAc:MeOH, 8:2) indicated the formation of the product (R_f = 0.0) with complete consumption of the starting material (R_f = 0.5). 20 μL of the crude were added to 1000 μL of MeOH. The sample was centrifuged and the absorbance of the solution measured by UV at 295 nm (A = 0.945).

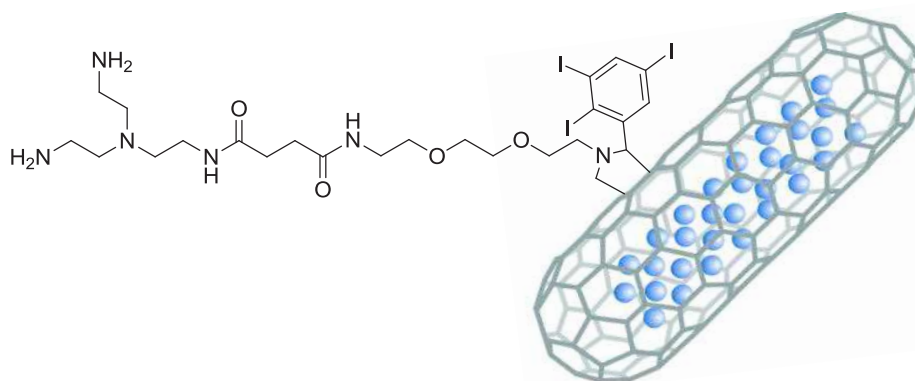
The functionalisation of the fullerene obtained was 0.742 mmol/g.

N-di(*tert*-butoxycarbonyl)-linker 2 NaI@SWNTs **5.22**



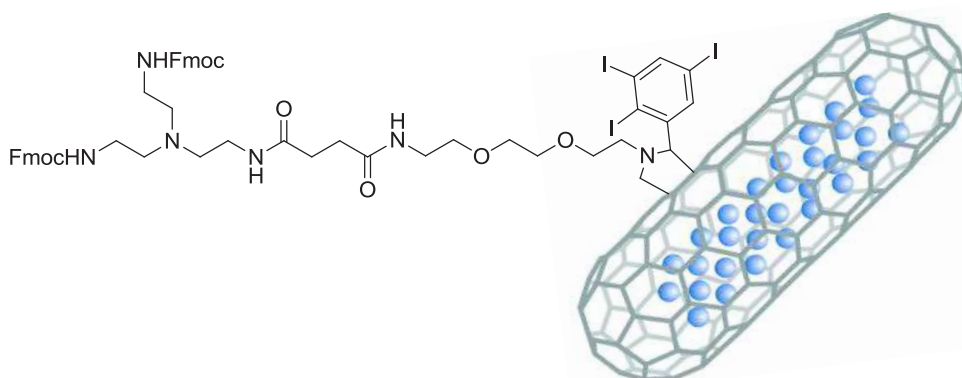
A suspension of **5.8** (6 mg) in DMF (dry, 3 mL) was sonicated for 2 min, before a solution of **5.18** (12.7 mg, 0.02 mmol, 1 eq.) and **5.4** (9.6 mg, 0.02 mmol, 1 eq.) in DMF (dry, 1 mL) was added. The reaction was refluxed at 130 °C for 4 days, then cooled to r.t. and filtered. The solid was washed with DMF, MeOH and dried under vacuum to afford the title compound **5.22** as a black solid (6.8 mg). **EA** (C,H,N) 65.66 ,1.13, 2.93.

(NH₂)₂-linker 2 NaI@SWNTs 5.23



A suspension of **5.22** (12.2 mg) in DCM (dry, 5 mL) was sonicated for 2 min, before TFA (2.5 mL) was added. The reaction was stirred at r.t. for 24 h, then evaporated, recovered with MeOH, filtered and washed with MeOH. The solid obtained was dried under vacuum to afford the title compound **5.23** as a black solid (9.46 mg). **EA** (C,H,N) 66.81,1.14, 3.07.

N-di(fluorenylmethyloxycarbonyl)-linker 2 Nal@SWNTs **5.24**



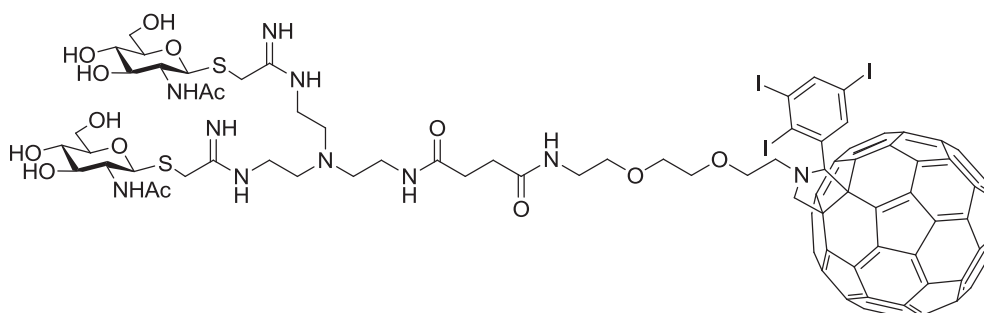
A solution of Fmoc chloride (1 mg, 3.86 μmol , 1 eq.) in DCM (0.5 mL) was added dropwise to a suspension of **5.23** (1 mg) in DCM (0.5 mL) at 0 °C. DIPEA (1.5 μL , 8.61 μmol , 2.2 eq.) was added and the reaction was stirred at r.t. for 16 h. The reaction mixture was centrifuged, the supernatant discarded and the solid recovered with MeOH then filtered and washed with MeOH. The solid obtained was dried under vacuum to afford the title compound **5.24** as a black solid (0.90 mg). **EA** (C,H,N) 67.57 , 0.98, 2.07.

Fmoc numbering on N-di(fluorenylmethyloxycarbonyl)-linker 2 Nal@SWNTs

5.24 (0.95 mg) were treated with DMSO:DMF:DBU (25:24:1 solution, 1 mL) at r.t. to cleave the Fmoc group. After 2 h, 20 μL of the crude were added to 1000 μL of MeOH. The sample was centrifuged and the absorbance of the solution measured by UV at 295 nm ($A = 0.065698$).

The functionalisation of the nanotubes obtained was 0.34 mmol/g.

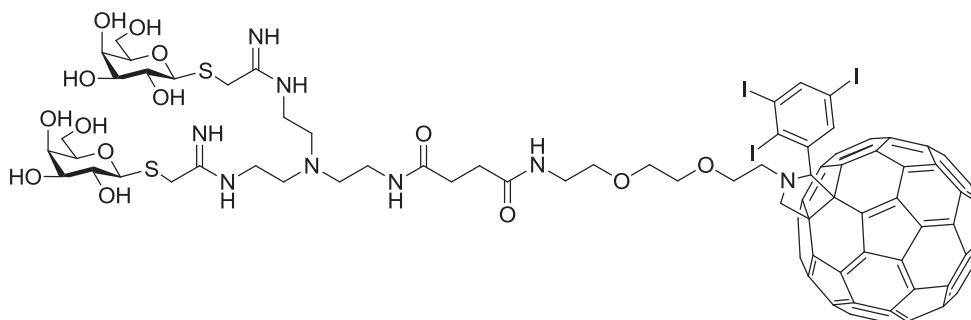
GlcNAc₂-linker 2 fullerene **5.25**



To a solution of **5.20** (14 mg, 9 μ mol, 1eq) in MeOH:DMSO (2:1, 1.5 mL) was added **2.80** (6 mg, 20 μ mol, 2.2 eq). The reaction was stirred at r.t. for 30 min, then filtered, washed with MeOH (100 mL) and Et₂O (100 mL) to afford the title compound **5.25** as a black solid (10.4 mg, 4.88 μ mol, 54%). **TLC** R_f = 0.10 (toluene:EtOAc, 4:1); **¹H NMR** (MeOD, 400 MHz) δ 8.31, 8.29 (2 x d, J = 1.9 Hz, 2 x 1 H, 2 x CH-Ar), 5.87 (s, 1 H, CH-pyrrolidine), 5.19 (d, J = 9.6 Hz, 1 H, CH₂-pyrrolidine), 4.69 (br, 2 x 1 H, 2 x H-1), 4.49 (d, J = 9.7 Hz, 1 H, CH₂-pyrrolidine), 3.98-3.92 (m, 4 H, 2 x H-2, 2 x H-4), 3.90-3.87 (m, 2 H, NCH₂CH₂O), 3.73-3.58 (m, 16 H, OCH₂CH₂O, 2 x SCH₂, 2 x H-3, 2 x H-5, 2 x H-6), 3.37 (t, J = 5.8 Hz, 2 H, OCH₂CH₂NHCOCH₂), 3.14-2.77 (m, 6 H, OCH₂CH₂N, OCH₂CH₂NHCOCH₂, NCH₂CH₂NHCOCH₂), 2.73 (br, 4 H, 2 x NCH₂CH₂NHC), 2.60 (br, 4 H, 2 x NCH₂CH₂NHC), 2.41 (m, 2 H, NCH₂CH₂NHCOCH₂), 2.27 (m, 4 H, COCH₂CH₂CO), 1.99, 1.91 (2 x s, 2 x 3 H, 2 x Ac); **¹³C NMR** (MeOD, 126 MHz) δ 173.52-137.89 (C-fullerene, C-Ar, CH-Ar, 2 x COCH₂, 2 x CNH, 2 x CONH), 114.42, 114.01, 97.34 (3 x Cl-Ar), 86.82 (CH-pyrrolidine), 84.33 (2 x C-1), 78.72 (2 x C-5), 73.54 (2 x C-3), 69.55, 69.43 (OCH₂CH₂O), 68.99

(OCH₂CH₂NHCOCH₂), 68.94 (OCH₂CH₂NHCOCH₂), 68.73 (OCH₂CH₂N), 68.51 (2 x C-2), 66.04 (CH₂-pyrrolidine), 62.17 (2 x SCH₂), 61.05 (2 x C-6), 52.70 (NCH₂CH₂NHCOCH₂), 51.89 (2 x NCH₂CH₂NHC), 50.56 (2 x C-4), 37.63 (OCH₂CH₂N), 36.87 (2 x NCH₂CH₂NHC), 36.12 (NCH₂CH₂NHCOCH₂), 31.59, 30.82 (COCH₂CH₂CO), 20.13 (2 x Ac); **MS** (LD⁺) m/z 2129.58 [M]; **IR** (ATR) 3298.76 (OH), 1717.88 (CO).

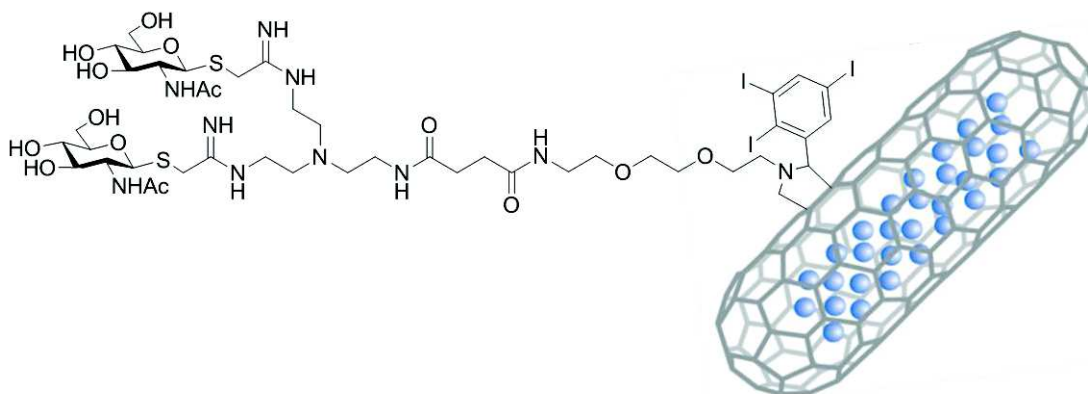
Gal₂-linker 2 fullerene **5.26**



To a solution of **5.20** (18 mg, 11 μmol, 1 eq.) in MeOH:DMSO (2:1, 1.5 mL) was added **2.77** (7.3 mg, 24.2 μmol, 2.2 eq). The reaction was stirred at r.t. for 30 min, then filtered, washed with MeOH (100 mL) and Et₂O (100 mL) to afford the title compound **5.26** as a black solid (14 mg, 4.88 μmol, 54%). **TLC** R_f = 0.10 (toluene:EtOAc, 4:1); **¹H NMR** (MeOD, 400 MHz) δ 8.30, 8.27 (2 x d, *J* = 1.9 Hz, 2 x 1 H, 2 x CH-Ar), 5.85 (s, 1 H, CH-pyrrolidine), 5.12 (d, *J* = 9.6 Hz, 1 H, CH₂-pyrrolidine), 4.54 (br, 2 x 1 H, 2 x H-1), 4.44 (d, *J* = 9.7 Hz, 1 H, CH₂-pyrrolidine), 3.99-3.87 (m, 2 H, 2 x H-4), 3.84-3.74 (m, 2 H, NCH₂CH₂O), 3.70-3.51 (m, 18 H, OCH₂CH₂O, 2 x SCH₂, 2 x H-2, 2 x H-3, 2 x H-5, 2 x H-6), 3.33 (t, *J* = 5.8 Hz, 2 H, OCH₂CH₂NHCOCH₂), 3.17-2.79 (m, 6 H, OCH₂CH₂N, OCH₂CH₂NHCOCH₂,

$\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 2.71 (br, 4 H, 2 x $\text{NCH}_2\text{CH}_2\text{NHC}$), 2.58 (br, 4 H, 2 x $\text{NCH}_2\text{CH}_2\text{NHC}$), 2.37 (m, 2 H, $\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 2.22 (m, 4 H, $\text{COCH}_2\text{CH}_2\text{CO}$); ^{13}C NMR (MeOD, 126 MHz) δ 177.20-134.83 (C-fullerene, C-Ar, CH-Ar, 2 x COCH_2 , 2 x CNH, 2 x CONH), 114.12, 114.00, 98.72 (3 x Cl-Ar), 88.03 (CH-pyrrolidine), 84.97 (2 x C-1), 79.22 (2 x C-5), 73.64 (2 x C-3), 69.53, 69.42 ($\text{OCH}_2\text{CH}_2\text{O}$), 68.96 (2 x C-4), 68.82 ($\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$), 68.81 ($\text{OCH}_2\text{CH}_2\text{NHCOCH}_2$), 68.74 ($\text{OCH}_2\text{CH}_2\text{N}$), 68.50 (2 x C-2), 66.33 (CH_2 -pyrrolidine), 62.12 (2 x SCH_2), 61.32 (2 x C-6), 53.54 ($\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 51.88 (2 x $\text{NCH}_2\text{CH}_2\text{NHC}$), 37.61 ($\text{OCH}_2\text{CH}_2\text{N}$), 36.39 (2 x $\text{NCH}_2\text{CH}_2\text{NHC}$), 35.99 ($\text{NCH}_2\text{CH}_2\text{NHCOCH}_2$), 31.58, 30.80 ($\text{COCH}_2\text{CH}_2\text{CO}$); MS (LD⁺) m/z 2047.47 [M]; IR (ATR) 3284.12 (OH), 1707.63 (CO).

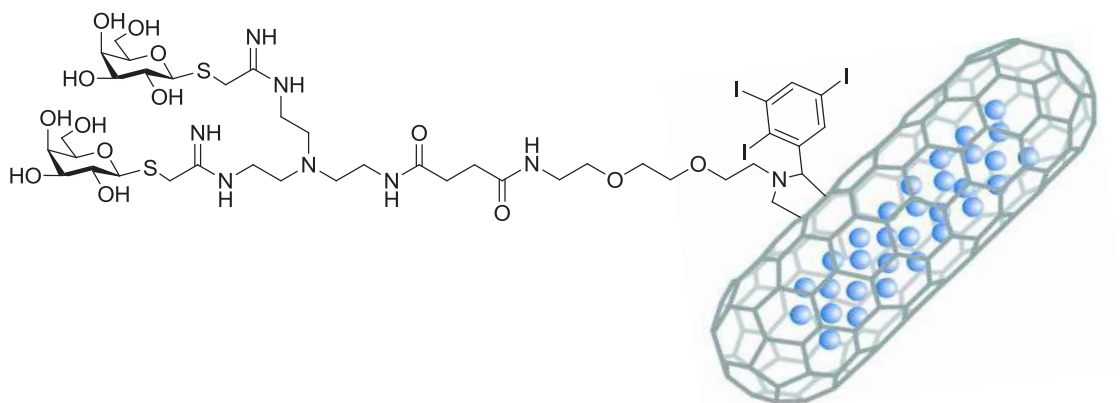
GlcNAc₂-linker 2 Nal@SWNTs 5.27



A solution of **2.80** (6.5 mg, 20 μmol , 1 eq.) in MeOH:DMSO (2:1, 1.5 mL) was added to a suspension of **5.23** (1 mg) in DCM (0.5 mL). The reaction was stirred at r.t. for 30 min. The reaction mixture was centrifuged, the supernatant discarded and the solid recovered with MeOH, filtered and washed with MeOH. The solid

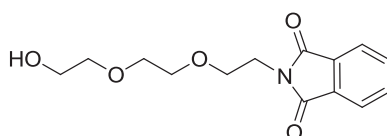
obtained was dried under vacuum to afford the title compound **5.27** as a black solid (0.88 mg). **EA** (C,H,N) 67.34, 1.08, 2.98.

Gal2-linker 2 Nal@SWNTs **5.28**



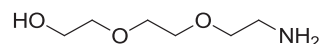
A solution of **2.77** (6.0 mg, 20 μ mol, 1 eq.) in MeOH:DMSO (2:1, 1.5 mL) was added to a suspension of **5.23** (1 mg) in DCM (0.5 mL). The reaction was stirred at r.t. for 30 min. The reaction mixture was centrifuged, the supernatant discarded and the solid recovered with MeOH, filtered and washed with MeOH. The solid obtained was dried under vacuum to afford the title compound **5.28** as a black solid (0.8 mg). **EA** (C,H,N) 67.32 ,1.11, 3.01.

2-[2-(2-hydroxyethoxy)ethoxy]methyl phthalimide **5.29**^[4]



To a suspension of potassium phthalimide (1.4 g, 7.6 mmol, 1.1 eq.) in DMF (7 mL) was added 2-[2-(2-chloroethoxy)ethoxy]ethanol (1 mL, 6.9 mmol, 1 eq.). The reaction mixture was stirred at 100 °C for 18 h. The reaction was filtered, evaporated, recovered with H₂O (150 mL) and extracted with DCM (3 x 150 mL). The combined organic layer was dried over MgSO₄, filtered and concentrated under vacuum, to afford the title compound **5.29** as a pale yellow oil (1.54 g, 5.8 mmol, 84%). **TLC** R_f = 0.2 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 500 MHz) δ 7.78 (dd, *J* = 5.4, 3.1 Hz, 2 H, 2 x CHC-arom), 7.66 (dd, *J* = 5.5, 3.0 Hz, 2 H, 2 x CHCH-arom), 3.84 (t, *J* = 5.8 Hz, 2 H, CH₂N), 3.69 (t, *J* = 5.8 Hz, 2 H, CH₂CH₂N), 3.63-3.57 (m, 4 H, OCH₂CH₂O), 3.55 (dd, *J* = 5.8, 2.8 Hz, 2 H, CH₂CH₂OH), 3.52-3.41 (m, 2 H, CH₂OH); **¹³C NMR** (CDCl₃, 101 MHz) δ 168.24 (2 x CO), 133.95 (2 x CHCH-arom), 131.95 (2 x C-arom), 123.16 (2 x CHC-arom), 72.44 (CH₂CH₂OH), 70.24, 69.89 (OCH₂CH₂O), 67.78 (CH₂CH₂N), 61.52 (CH₂OH), 37.10 (CH₂N); **MS** (FI⁺) *m/z* = 280.29 [M+H]; **IR** (ATR) 3382.16 (OH), 2871.43 (CH), 1705.74 (CO), 873.81, 794.33, 717.65 (arom).

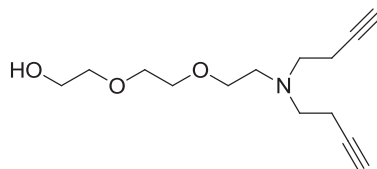
2-[2-(2-aminoethoxy)ethoxy]ethanol **5.30**^[3]



To a solution of **5.29** (500 mg, 1.8 mmol, 1 eq.) in EtOH (1.8 mL) was added hydrazine monohydrate (1.2 mL, 24 mmol, 13 eq.) and the mixture was stirred at 100 °C. After 1h a white ppt appeared. The reaction mixture was recovered with EtOH, filtered and the solid washed with EtOH. The solution was evaporated

under vacuum to afford the title compound **5.30** as a white solid (290.5 mg, 1.8 mmol, quant.). **TLC** R_f = 0.0 (PE:EtOAc, 1:1); **^1H NMR** (CDCl_3 , 250 MHz) δ 4.22 (br, 3 H, OH, NH_2), 3.79-3.68 (m, 2 H, CH_2NH_2), 3.68-3.45 (m, 8 H, $\text{CH}_2\text{OCH}_2\text{CH}_2\text{OCH}_2$), 2.91 (t, J = 5.1 Hz, 2 H, CH_2OH); **^{13}C NMR** (CDCl_3 , 63 MHz) δ 73.19 ($\text{CH}_2\text{CH}_2\text{OH}$), 72.43 ($\text{CH}_2\text{CH}_2\text{NH}$), 70.67, 70.50 ($\text{OCH}_2\text{CH}_2\text{O}$), 61.63 (CH_2OH), 41.48 (CH_2NH_2); **MS** (ESI^+) m/z = 150.20 [$\text{M}+\text{H}$]; **IR** (ATR) 3278.63 (OH), 2869.71 (CH).

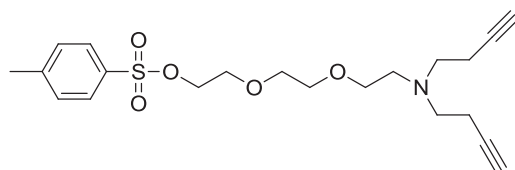
2-[2-(2-(dibut-3-ynyl aminoethoxy)ethoxy)]ethanol **5.31**



To a solution of **5.30** (50 mg, 0.33 mmol, 1 eq.) in ACN (dry, 2 mL) were added KI (22 mg, 0.13 mmol, 0.4 eq.), Cs_2CO_3 (437 mg, 1.34 mmol, 4 eq.) and 4-bromobutyne (69 μL , 0.74 mmol, 2.2 eq.). The mixture was sealed in a vial and heated by MW at 85 $^\circ\text{C}$ (250 W) for 1 h. A second portion of 4-bromobutyne (69 μL , 0.74 mmol, 2.2 eq.) was added before repeating the heating treatment. This was repeated for a third and a fourth addition of 4-bromobutyne (69 μL x 2, 0.74 mmol x 2, 2.2 eq. x 2). Finally, the reaction was allowed to cool to r.t.. The reaction was filtered, evaporated and purified by flash column chromatography on silica (EtOAc:MeOH, 10:0.25), to afford the title compound **5.31** as a colorless oil (45.2 mg, 0.18 mmol, 54%). **TLC** R_f = 0.8 (DCM:MeOH, 10:1); **^1H NMR** (CDCl_3 , 400 MHz) δ 3.72-3.67 (m, 2 H, CH_2OH), 3.66-3.61 (m, 2 H, $\text{OCH}_2\text{CH}_2\text{O}$), 3.61-3.56 (m, 4 H, $\text{OCH}_2\text{CH}_2\text{O}$, $\text{CH}_2\text{CH}_2\text{OH}$), 3.53 (t, J = 6.0 Hz, 2 H, $\text{CH}_2\text{CH}_2\text{N}$), 2.74 (m,

6 H, CH₂CH₂N, 2 x NCH₂CH₂C), 2.40-2.24 (m, 4 H, 2 x CH₂C), 1.94 (t, *J* = 2.7 Hz, 2 H, 2 x CCH); ¹³C NMR (CDCl₃, 101 MHz) δ 82.81 (2 x CCH), 72.54 (OCH₂CH₂O), 70.52 (CH₂CH₂OH), 70.39 (OCH₂CH₂O), 69.88 (CH₂CH₂N), 69.10 (2 x CCH), 61.72 (CH₂OH), 53.24, 53.21 (CH₂CH₂N, 2 x NCH₂CH₂C), 17.26 (2 x CH₂C); MS (ESI⁺) *m/z* = 254.34 [M+H]; HRMS (ESI⁺) found 254.1759 [M+H], C₁₄H₂₄NO₃ requires 254.1751; IR (ATR) 3283.98 (OH), 2923.72 (CH), 2358.95 (CCH-alkyne).

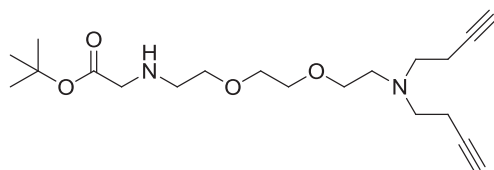
2-[2-(2-(dibut-3-ynyl aminoethoxy)ethoxy)]ethyl 4-methylbenzenesulfonate 5.32



To a solution of **5.31** (136.4 mg, 0.54 mmol, 1 eq.) in DCM (5 mL), were added pyridine (90 µL, 1.08 mmol, 2 eq.) and tosyl chloride (250 mg, 1.08 mmol, 2 eq.). The reaction was stirred at r.t. for 16 h, then recovered with DCM (100 mL), washed with HCl 1 M (2 x 150 mL) and brine (150 mL). The organic layer was dried over MgSO₄ and evaporated under vacuum. The crude was purified by flash column chromatography over silica (PE:EtOAc, 8:2) to afford the title compound **5.32** as a yellow oil (111.3 mg, 0.27 mmol, 51%). TLC *R_f* = 0.2 (DCM:MeOH, 10:1); ¹H NMR (CDCl₃, 400 MHz) δ 7.78 (d, *J* = 8.3 Hz, 2 H, 2 x CHCSO₂-arom), 7.32 (d, *J* = 8.1 Hz, 2 H, 2 x CHCCH₃-arom), 4.17-4.09 (m, 2 H, SO₂OCH₂), 3.71-3.60 (m, 2 H, SO₂OCH₂CH₂), 3.55-3.48 (m, 6 H, OCH₂CH₂O, CH₂CH₂N), 2.72 (m, 6 H, CH₂CH₂N, 2 x NCH₂CH₂C), 2.43 (s, 3 H, CH₃), 2.29 (td, *J* = 7.5, 2.7 Hz,

4 H, 2 x CH₂C), 1.94 (t, *J* = 2.7 Hz, 2 H, 2 x CCH); **¹³C NMR** (CDCl₃, 101 MHz) δ 144.81 (CSO₂), 129.83 (2 x CHCCH₃-arom), 127.98 (2 x CHCSO₂-arom), 82.84 (CCH₃-arom), 70.76, 70.43 (OCH₂CH₂O), 69.95 (CH₂CH₂N), 69.26 (CH₂OSO₂), 69.09 (2 x CCH), 68.72 (CH₂CH₂OSO₂), 53.22, 53.17 (CH₂CH₂N, 2 x NCH₂CH₂C), 21.66 (CCH₃), 17.31 (2 x CH₂CCH); **MS** (ESI⁺) *m/z* = 430.52 [M+Na]; **HRMS** (ESI⁺) found 430.1646 [M+Na], C₂₁H₂₉NO₅SNa requires 430.1664; **IR** (ATR) 3289.31 (CCH-alkyne), 2868.28 (CH), 2235.77 (CCH-alkyne), 826.69 (arom).

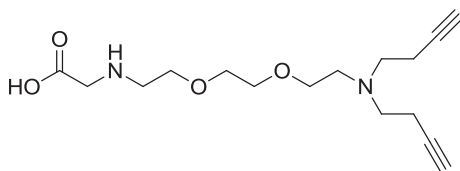
N-2-(2-(2-(dibut-3-ynyl aminoethoxy)ethoxy)ethyl)glycine (*tert*-butyl)ester
5.33



To a solution of **5.32** (111.3 mg, 0.27 mmol, 1.1 eq.) in ACN (dry, 9 mL), were added glycine *tert*-butyl ester hydrochloride (42 mg, 0.25 mmol, 1 eq.) and Cs₂CO₃ (202 mg, 0.62 mmol, 2.5 eq.). The reaction was stirred at r.t. for 24 h, then filtered and evaporated under vacuum. The crude was purified by flash column chromatography over silica (DCM:MeOH, 10:0.25) to afford the title compound **5.33** as a yellow oil (6.6 mg, 0.018 mmol, 8%). **TLC** *R_f* = 0.8 (DCM:MeOH, 10:1); **¹H NMR** (CDCl₃, 400 MHz) δ 3.66-3.51 (m, 8 H, CH₂OCH₂CH₂OCH₂), 3.34 (s, 2 H, NHCH₂CO), 2.78 (td, *J* = 13.4, 6.0 Hz, 8 H, OCH₂CH₂N, 2 x NCH₂CH₂CCH, OCH₂CH₂NH), 2.34 (td, *J* = 7.5, 2.6 Hz, 4 H, 2 x CH₂CCH), 2.10 (br, 1 H, NH), 1.96 (t, *J* = 2.7 Hz, 2 H, CCH), 1.46 (s, 9 H, ^{*t*}Bu);

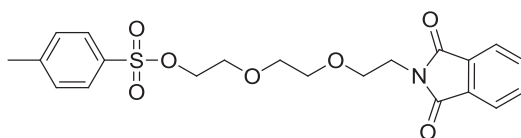
¹³C NMR (CDCl₃, 101 MHz) δ 171.14 (COO), 82.70 (C(CH₃)₃), 81.20 (2 x CCH), 70.60, 70.47 (OCH₂CH₂O), 70.39 (OCH₂CH₂N), 69.85 (OCH₂CH₂NH), 69.07 (2 x CCH), 53.25, 53.21 (OCH₂CH₂N, 2 x NCH₂CH₂CCH), 51.59 (NHCH₂CO), 48.73 (OCH₂CH₂NH), 28.13 (C(CH₃)₃), 17.31 (CH₂CCH); **MS** (ESI⁺) m/z = 367.50 [M+H]; **HRMS** (ESI⁺) found 367.2587 [M+H], C₂₀H₃₅N₂O₄ requires 367.2591; **IR** (ATR) 3277.53 (CCH-alkyne), 2890.13 (CH), 2223.89 (CCH-alkyne), 1735.78 (CO).

Linker 3 5.34



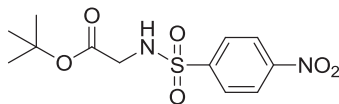
To **5.33** (6.6 mg, 0.018 mmol, 1 eq.) was added DCM:TFA (1:1, 4.8 mL) and the reaction stirred at r.t. for 6 h before being evaporated under vacuum to afford the title compound **5.34** as a yellow oil (7.8 mg, 0.018 mmol, quant). **TLC** R_f = 0.0 (DCM:MeOH, 10:1); **¹H NMR** (MeOD, 400 MHz) δ 3.88 (br, 4 H, CH₂OCH₂CH₂OCH₂), 3.79 (br, 2 H, NHCH₂CO), 3.72 (br, 4 H, OCH₂CH₂O), 3.59-3.48 (m, 6 H, OCH₂CH₂N, 2 x NCH₂CH₂CCH), 2.79 (dd, *J* = 6.9, 4.6 Hz, 4 H, 2 x CH₂CCH), 2.63 (t, *J* = 2.3 Hz, 2 H, 2 x CCH); **¹³C NMR** (MeOD, 101 MHz) δ 78.07 (2 x CCH), 72.13 (2 x CCH), 69.96 (OCH₂CH₂O), 65.58 (NHCH₂CO), 64.39 (OCH₂CH₂N), 52.82 (2 x NCH₂CH₂CCH), 51.87 (OCH₂CH₂N), 46.82 (OCH₂CH₂NH), 45.59 (OCH₂CH₂NH), 13.76 (2 x CH₂CCH); **MS** (ESI⁺) m/z = 311.40 [M+H]; **HRMS** (ESI⁺) found 311.1972 [M+H], C₁₆H₂₇N₂O₄ requires 311.1965; **IR** (ATR) 3324.42 (OH), 1671.88 (CO).

2-[2-(2-(phthalimidoethoxy)ethoxy)]ethyl 4-methylbenzenesulfonate 5.35^[5]



To a solution of **5.29** (1.52 g, 1.52 mmol, 1 eq.) in DCM (3 mL) was added pyridine (245 μ L, 3.03 mmol, 2 eq.) and tosyl chloride (576 mg, 3.03 mmol, 2 eq.). The reaction mixture was stirred at r.t for 18 h before being diluted with DCM (100 mL), washed with HCl 0.1 N (2 x 200 mL) and brine (200 mL). The organic layer was dried over MgSO_4 , evaporated under vacuum and purified by flash column chromatography over silica (EtOAc:hexane, 1:3 to 1:1), to afford the title compound **5.35** as a colourless oil (410 mg, 0.95 mmol, 62%). **TLC** R_f = 0.75 (EtOAc:hexane, 1:1); **$^1\text{H NMR}$** (CDCl_3 , 400 MHz) δ 7.80 (dd, J = 5.5, 3.0 Hz, 2 H, 2 x CHC-NPht), 7.76, 7.74 (2 x s, 2 x 1 H, 2 x CHCSO_2 -arom), 7.69 (dd, J = 5.5, 3.0 Hz, 1 H, 2 x CHCHC-NPht), 7.31 (d, J = 8.0 Hz, 2 H, 2 x CHCCH_3 -arom), 4.07 (dd, J = 5.4, 4.3 Hz, 2 H, CH_2N), 3.85 (t, J = 5.8 Hz, 2 H, CH_2OSO_2), 3.67 (t, J = 5.8 Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{OSO}_2$), 3.60 (dd, J = 5.4, 4.3 Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{N}$), 3.54, 3.50 (2 x dd, J = 5.6, 2.5 Hz, 2 x 2 H, $\text{OCH}_2\text{CH}_2\text{O}$), 2.41 (s, 3 H, CH_3); **$^{13}\text{C NMR}$** (CDCl_3 , 101 MHz) δ 168.21 (2 x CO), 144.77 (CSO_2), 133.98 (2 x CHCHC-NPht), 132.92 (CCH_3), 132.03 (2 x C-NPht), 129.79 (2 x CHCCH₃), 127.92 (2 x CHCSO_2), 123.20 (2 x CHC-NPht), 70.62, 69.98 ($\text{OCH}_2\text{CH}_2\text{O}$), 69.20 (CH_2N), 68.63 ($\text{OCH}_2\text{CH}_2\text{N}$), 67.92 ($\text{OCH}_2\text{CH}_2\text{OSO}_2$), 37.17 ($\text{OCH}_2\text{CH}_2\text{OSO}_2$), 21.61 (CH_3); **MS** (FI^+) m/z = 433.47 [M]; **IR** (ATR) 1707.55 (CO), 816.24, 754.14 (arom).

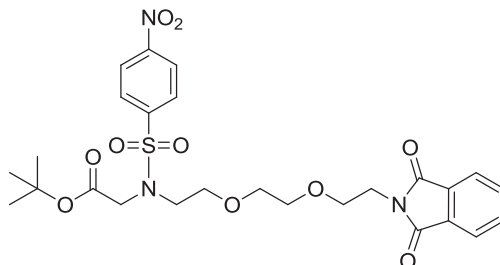
N-nosyl glycine (*tert*-butyl)ester **5.36**^[6]



To a solution of glycine *tert*-butyl ester hydrochloride (1 g, 6 mmol, 1 eq.) in DCM (50 mL) were added nosyl chloride (1.59 g, 7.2 mmol, 1.2 eq.) and collidine (932 μ L, 7.2 mmol, 1.2 eq.) in DCM (10 mL). The mixture was stirred for 22 h at r.t. before adding more nosyl chloride (800 mg, 3.6 mmol, 0.6 eq.) and collidine (460 μ L, 3.6 mmol, 0.6 eq.) in DCM (3 mL). After a further 11 h the reaction mixture was evaporated, recovered with DCM (150 mL) and washed with HCl 1N (200 mL) and brine (200 mL). The organic layer was dried over MgSO_4 and the solvent was removed under vacuum. The residue was purified by flash column chromatography over silica (PE:DCM, 8:2) to afford the title compound **5.36** as a pink solid (1.54 g, 4.9 mmol, 82%). **TLC** R_f = 0.2 (DCM:MeOH, 10:1); **^1H NMR** (CDCl_3 , 400 MHz) δ 8.01-7.97 (m, 1 H, $\text{CHCNO}_2\text{-arom}$), 7.84-7.79 (m, 1 H, $\text{CHCSO}_2\text{-arom}$), 7.73-7.60 (m, 2 H, CHCHCHCH-arom), 6.01 (t, J = 5.7 Hz, 1 H, NH), 3.81 (d, J = 5.8 Hz, 2 H, CH_2), 1.22 (s, 9 H, ^tBu); **^{13}C NMR** (CDCl_3 , 101 MHz) δ 167.49 (CO), 147.76 (CNO_2), 133.91, 133.56 (CHCHCHCH-arom), 133.04 (CSO_2), 130.55 ($\text{CHCNO}_2\text{-arom}$), 125.57 ($\text{CHCSO}_2\text{-arom}$), 82.88 ($\text{C}(\text{CH}_3)_3$), 45.53 (CH_2), 27.69 ($\text{C}(\text{CH}_3)_3$); **MS** (FI) m/z = 315.07 [M-H]; **IR** (ATR) 3338.59 (NH), 2982.66 (CH), 1726.82 (CO), 734.08 (arom).

N-nosyl-N-2-(2-(2-(phthalimidoethoxy)ethoxy)ethyl)glycine (*tert*-butyl)ester

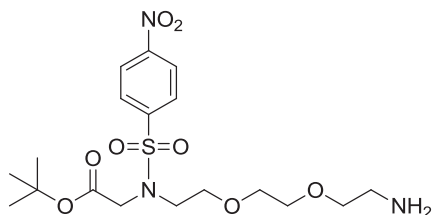
5.37



To a solution of **5.35** (151 mg, 0.34 mmol, 1.1 eq.) in ACN (dry, 6 mL), were added **5.36** (100 mg, 0.31 mmol, 1 eq.) and Cs₂CO₃ (258 mg, 0.79 mmol, 2.5 eq.). The reaction was initially stirred at r.t. for 16 h, then refluxed at 85 °C for 24 h, filtered and evaporated under vacuum. The crude was purified by flash column chromatography over silica (EtOAc:PE, 9:1) to afford the title compound **5.37** as a yellow oil (45.6 mg, 0.08 mmol, 25%). **TLC** R_f = 0.15 (PE:EtOAc, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 7.98 (d, *J* = 10.5 Hz, 1 H, CHCNO₂), 7.86-7.76 (m, 2 H, CHCHCHCH-NPht), 7.76-7.61 (m, 3 H, CHCHCHCH-NPht, CHCSO₂), 7.29 (d, *J* = 9.7 Hz, 2 H, CHCHCHCH-NS), 5.11 (t, *J* = 5.8 Hz, 2 H, CH₂CO), 3.88 (dt, *J* = 11.6, 5.8 Hz, 2 H, OCH₂CH₂NSO₂), 3.80 (t, *J* = 5.9 Hz, 2 H, OCH₂CH₂NPht), 3.75-3.42 (m, 8 H, CH₂OCH₂CH₂OCH₂), 1.65 (s, 9 H, ^{*t*}Bu); **¹³C NMR** (CDCl₃, 101 MHz) δ 168.19 (COO), 157.69 (2 x CON), 133.87 (2 x CHC-NPht), 132.12 (CNO₂), 129.42 (CSO₂), 127.25 (CHCHCSO₂), 126.36 (CHCHCNO₂), 123.20 (2 x CHCH-NPht), 121.49 (CHCNO₂), 119.64 (2 x C-NPht), 113.44 (CHCSO₂), 82.72 (C(CH₃)₃), 70.35, 70.03 (OCH₂CH₂OCH₂CH₂NPht), 67.88 (OCH₂CH₂NSO₂), 67.82 (OCH₂CH₂NSO₂), 44.37 (CH₂COO), 37.19 (OCH₂CH₂NPht), 28.46

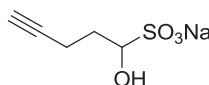
(C(CH₃)₃); **MS** (ESI⁺) m/z = 578.61 [M+H]; **HRMS** (ESI⁺) found 578.1809 [M+H], C₂₆H₃₂N₃O₁₀S requires 578.1803; **IR** (ATR) 1735.15 (COO), 1690.91 (CONH).

N-nosyl-N-2-(2-(2-(aminoethoxy)ethoxy)ethyl)glycine (tert-butyl)ester 5.38



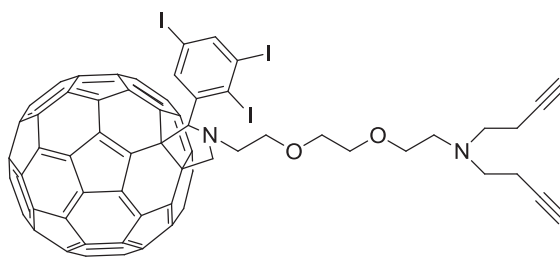
To a solution of **5.37** (45.6 mg, 0.08 mmol, 1 eq.) in EtOH (1 mL), was added hydrazine (50 μ L, 1.03 mmol, 13 eq.). The reaction was refluxed at 85 °C for 3 h, then filtered and evaporated under vacuum to afford the title compound **5.38** as yellow solid (40 mg, 0.08 mmol, quant.). **TLC** R_f = 0.1 (PE:EtOAc, 9:1); **¹H NMR** (MeOD, 400 MHz) δ 8.20 (m, 2 H, CHCNO₂, CHCSO₂), 7.79 (m, 2 H, CHCHCHCH-arom), 5.16 (br, 2 H, CH₂CO), 3.95 (br, 2 H, OCH₂CH₂N), 3.73-3.44 (m, 8 H, CH₂OCH₂CH₂OCH₂), 2.99-2.85 (m, 2 H, CH₂NH₂), 1.69 (s, 9 H, ^tBu); **¹³C NMR** (MeOD, 101 MHz) δ 159.61 (CO), 131.51 (CHCHCSO₂), 129.32 (CNO₂), 127.46 (CSO₂), 125.53 (CHCHCNO₂), 121.28 (CHCNO₂), 112.19 (CHCSO₂), 83.23 (C(CH₃)₃), 72.22 (OCH₂CH₂NH₂), 70.05, 69.91 (OCH₂CH₂O), 67.99, 67.63 (OCH₂CH₂N), 44.66 (CH₂CO), 39.56 (CH₂NH₂), 27.28 (C(CH₃)₃); **MS** (ESI⁺) m/z = 448.51 [M+H]; **HRMS** (ESI⁺) found 448.1741 [M+H], C₁₈H₃₀N₃O₈S requires 448.1748; **IR** (ATR) 3345.76, 3344.18 (NH₂), 1730.75 (CO).

Sodium 1-hydroxypent-4-yne-1-sulphonate **5.39**



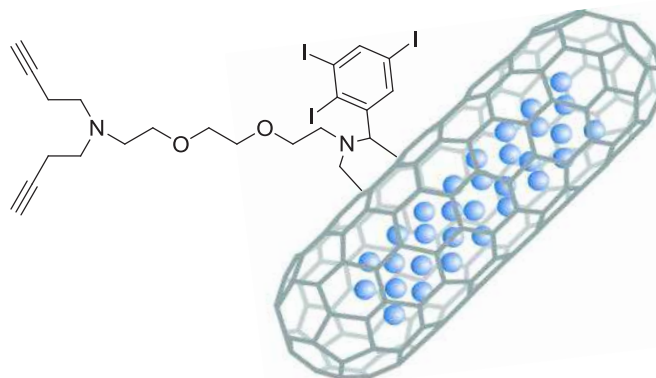
To a suspension of PCC (1.7 g, 7.89 mmol, 1.2 eq.) in DCM (dry, 8 mL) under argon was added 4-pentynol (500 μ L, 6.57 mmol, 1 eq.) in DCM (1 mL), dropwise, at 0 °C. The mixture was stirred at r.t for 3 h before more PCC (500 mg, 2.32 mmol, 0.35 eq.) and EtOH (1 mL) were added. After other 5 h the reaction mixture was filtered through celite, the solid triturated with EtOH (4 x 6 mL) then the solutions were collected together, shaken with silica, filtered, and the silica washed with EtOAc (50 mL). To the solution was added NaHSO₃ (750 mg, 6.25 mmol, 0.95 eq.) in H₂O (2.5 mL), dropwise. Some EtOH was added to help the precipitation. The reaction was stirred for 16 h at r.t., before the solvent was removed under vacuum. The residue was triturated with EtOH and then recrystallised from hot EtOH/H₂O, to afford the title compound **5.39** as a white solid (1.06 g, 5.69 mmol, 87%). **TLC** R_f = 0.0 (PE:EtOAc, 9:1); **¹H NMR** (D₂O, 400 MHz) δ 4.72 (m, 1 H, CHOH, CHCSO₂), 2.83 (s, 1 H, CCH), 2.17 (m, 2 H, CHCH₂CH₂C), 2.03 (br, 2 H, CHCH₂CH₂C); **¹³C NMR** (D₂O, 101 MHz) δ 97.01 (CHOH), 84.70 (CCH), 71.31 (CCH), 29.57 (CHCH₂CH₂C), 13.36 (CHCH₂CH₂C); **MS** (ESI⁻) m/z = 185.15 [M-H]; **HRMS** (ESI⁻) found 184.9888 [M-H], C₅H₆NaO₄S requires 184.9890; **IR** (ATR) 3375.42 (OH), 2201.77 (CCH-alkyne).

Linker 3 fullerene **5.40**



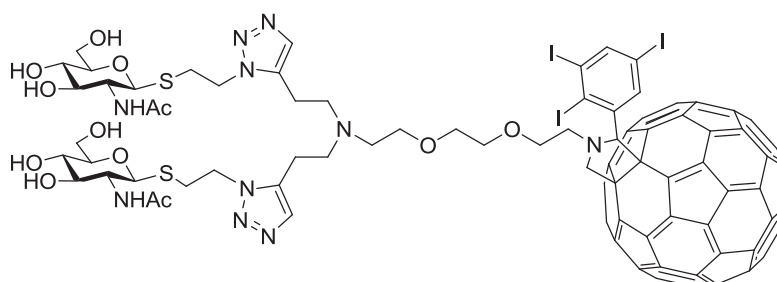
To a solution of C₆₀ (18 mg, 0.025 mmol, 1 eq.) in toluene (10 mL) were added **5.34** (7.8 mg, 0.025 mmol, 1 eq.) and **5.4** (12.2 mg, 0.025 mmol, 1 eq.). The reaction was refluxed at 120 °C for 3 h, then cooled to r.t. and evaporated under vacuum. The crude was purified by flash column chromatography over silica (toluene:EtOAc, 10:0.5 to 10:1) to afford the title compound **5.40** as a brown solid (8.2 mg, 5.64 μmol, 23%). **TLC** R_f = 0.5 (EtOAc:toluene, 1:10); **¹H NMR** (CDCl₃, 250 MHz) δ 8.41, 8.22 (2 x s, 2 x 1 H, 2 x CH-arom), 5.83 (s, 1 H, CH-pyrrolidine), 5.18, 4.44 (2 x d, *J* = 9.6 Hz, 2 x 1 H, CH₂-pyrrolidine), 3.96 (t, *J* = 5.4 Hz, 2 H, OCH₂CH₂Npyrrolidine), 3.78 (br, 4 H, OCH₂CH₂O), 3.70-3.57 (m, 2 H, OCH₂CH₂Npyrrolidine), 3.26 (br, 2 H, OCH₂CH₂NCH₂), 2.92 (br, 2 H, OCH₂CH₂NCH₂), 2.80 (br, 2 x 2 H, 2 x CH₂NCH₂CH₂C), 2.36 (s, 2 x 1 H, 2 x CCH), 1.57 (br, 2 x 2 H, 2 x CH₂NCH₂CH₂C); **¹³C NMR** (CDCl₃, 126 MHz) δ 156.37-125.27 (C-fullerene, C-arom, CH-arom), 115.28, 112.46, 94.94 (3 x Cl-arom), 86.17 (CH-pyrrolidine), 75.70 (2 x CCH), 70.67 (OCH₂CH₂Npyrrolidine), 69.98 (OCH₂CH₂O), 69.31 (OCH₂CH₂Npyrrolidine), 66.93 (CH₂-pyrrolidine), 53.16 (OCH₂CH₂NCH₂), 51.76 (OCH₂CH₂NCH₂), 41.72 (2 x CCH), 29.68 (2 x CH₂NCH₂CH₂C), 21.44 (2 x CH₂NCH₂CH₂C); **MS** (LD⁺) *m/z* = 1453.65 [M+H]; **IR** (ATR) 2210.46 (CCH-alkyne).

Linker 3 Nal@SWNTs 5.41



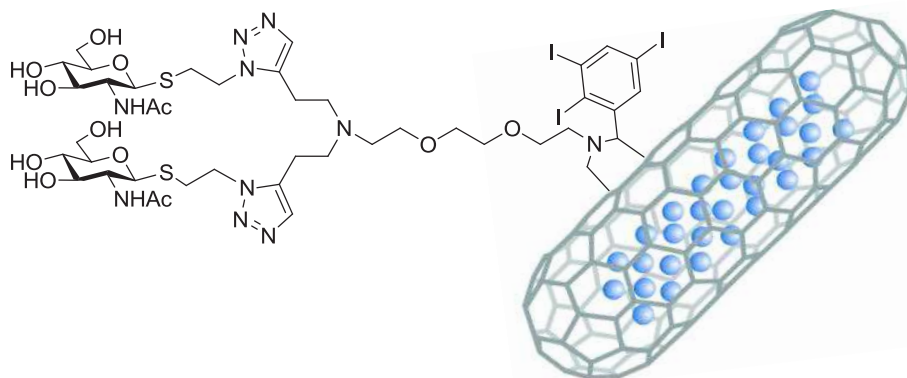
A suspension of Nal@SWNTs (6 mg) in DMF (dry, 3 mL) was sonicated for 2 min, before a solution of **5.34** (6.2 mg, 0.02 mmol, 1 eq.) and **5.4** (9.6 mg, 0.02 mmol, 1 eq.) in DMF (dry, 1 mL) was added. The reaction was refluxed at 130 °C for 4 days, then cooled to r.t. and filtered. The solid was washed with DMF, MeOH and dried under vacuum to afford the title compound **5.41** as a black solid (0.9 mg). **EA** (C,H,N) 66.52 ,0.87, 1.09.

GlcNAc₂-linker 3 fullerene 5.42



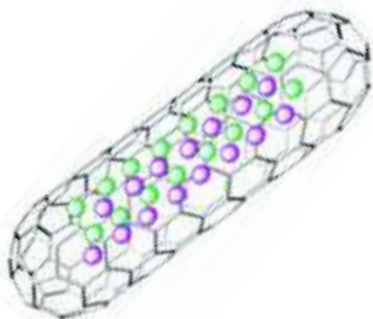
To a solution of **5.40** (14.5 mg, 10 μ mol, 1 eq.) in MeOH:DMSO (2:1, 1.5 mL) was added **3.7** (6.7 mg, 22 μ mol, 2.2 eq) and a solution of CuBr and TTMA pre-mixed in MeOH (CuBr: 0.43 mg, 3 μ mol, 0.3 eq.; TTMA: 3.1 mg, 6 μ mol, 0.6 eq.; MeOH, 100 μ L). The reaction was stirred at r.t. for 16 h, then filtered and washed with MeOH (100 mL) to afford the title compound **5.42** as a black solid (17.9 mg, 8.67 μ mol, 87%). **TLC** R_f = 0.15 (toluene:EtOAc, 4:1); **^1H NMR** (MeOD, 400 MHz) δ 8.37, 8.28 (2 x d, J = 1.9 Hz, 2 x 1 H, 2 x CH-Ar), 7.30 (2 x CH-triazole), 5.81 (s, 1 H, CH-pyrrolidine), 5.11 (d, J = 9.6 Hz, 1 H, CH₂-pyrrolidine), 4.66 (br, 2 x 1 H, 2 x H-1), 4.44 (d, J = 9.7 Hz, 1 H, CH₂-pyrrolidine), 4.08-3.95 (m, 4 x 1 H, 2 x H-2, 2 x H-4), 3.93-3.89 (m, 2 H, OCH₂CH₂NCH₂), 3.73-3.58 (m, 20 H, OCH₂CH₂O, 2 x SCH₂CH₂, 2 x H-3, 2 x H-5, 2 x H-6), 3.35 (t, J = 5.8 Hz, 2 H, OCH₂CH₂NCH₂), 3.11-2.97 (m, 4 H, OCH₂CH₂Npyrrolidine, OCH₂CH₂Npyrrolidine), 2.63 (br, 4 H, 2 x CH₂NCH₂CH₂C), 2.59 (br, 4 H, 2 x CH₂NCH₂CH₂C), 2.01, 1.98 (2 x s, 2 x 3 H, 2 x Ac); **^{13}C NMR** (MeOD, 126 MHz) δ 177.82-136.71 (C-fullerene, C-Ar, CH-Ar, 2 x CONH, 2 x C-triazole, 2 x CH-triazole), 116.87, 114.99, 101.34 (3 x Cl-Ar), 86.80 (CH-pyrrolidine), 84.31 (2 x C-1), 78.52 (2 x C-5), 73.44 (2 x C-3), 68.99, 68.71 (OCH₂CH₂O), 68.11 (OCH₂CH₂NCH₂), 68.04 (OCH₂CH₂NCH₂), 67.74 (OCH₂CH₂Npyrrolidine), 67.50 (2 x C-2), 67.04 (CH₂-pyrrolidine), 64.94 (2 x SCH₂CH₂), 64.33 (2 x SCH₂CH₂), 62.05 (2 x C-6), 55.73 (OCH₂CH₂Npyrrolidine), 51.88 (2 x CH₂NCH₂CH₂C), 50.56 (C-4), 36.87 (2 x CH₂NCH₂CH₂C), 20.07 (2 x Ac); **MS** (LD⁺) m/z 2065.49 [M]; **IR** (ATR) 3287.44 (OH), 1705.61 (CO).

GlcNAc₂-linker **3** NaI@SWNTs **5.43**



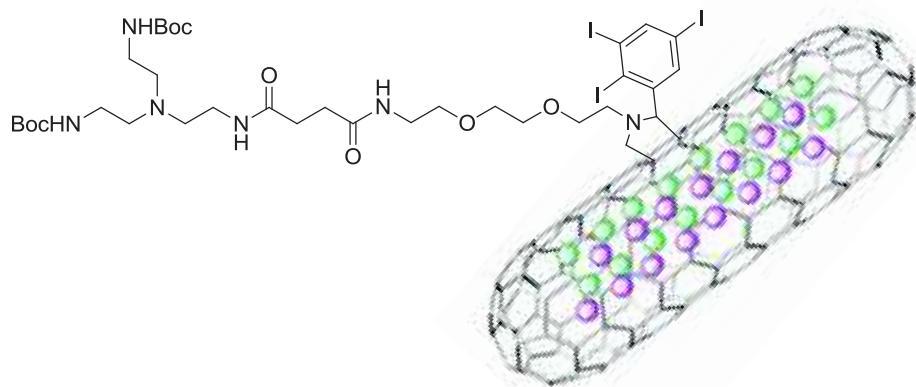
To a solution of **5.41** (6.7 mg, 22 μmol , 1 eq.) in MeOH:DMSO (2:1, 1.5 mL) were added **3.7** (1 mg) and a solution of CuBr and TTMA pre-mixed in MeOH (CuBr: 0.43 mg, 3 μmol , 0.3 eq.; TTMA: 3.1 mg, 6 μmol , 0.6 eq.; MeOH, 100 μL). The reaction was stirred at r.t. for 16 h. The reaction mixture was centrifuged, the supernatant discarded and the solid recovered with MeOH, filtered and washed with MeOH. The solid obtained was dried under vacuum to afford the title compound **5.43** as a black solid (0.98 mg). **EA** (C,H,N) 67.65 , 1.92, 3.17.

Na¹²⁵I@SWNTs **5.44**



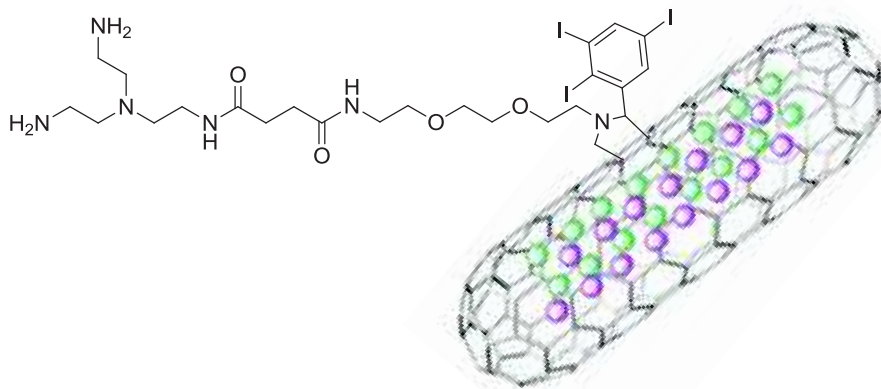
Steam-purified SWNTs (1 mg) were loaded in a quartz ampoule with Na¹²⁵I solution (10⁻⁵ M, 100 μ L, 20 MBq), dried and sealed under high vacuum with an oxygen-propane flame. The ampoule was heated in a furnace up to 900 °C using the same ramp described for **5.8**. The sample was recovered from the ampoule with H₂O (1 mL), sonicated and filtered. Final activity = 9 MBq (RCY = 45%).

N-di(*tert*-butoxycarbonyl)-linker 2 Na¹²⁵I@SWNTs **5.45**



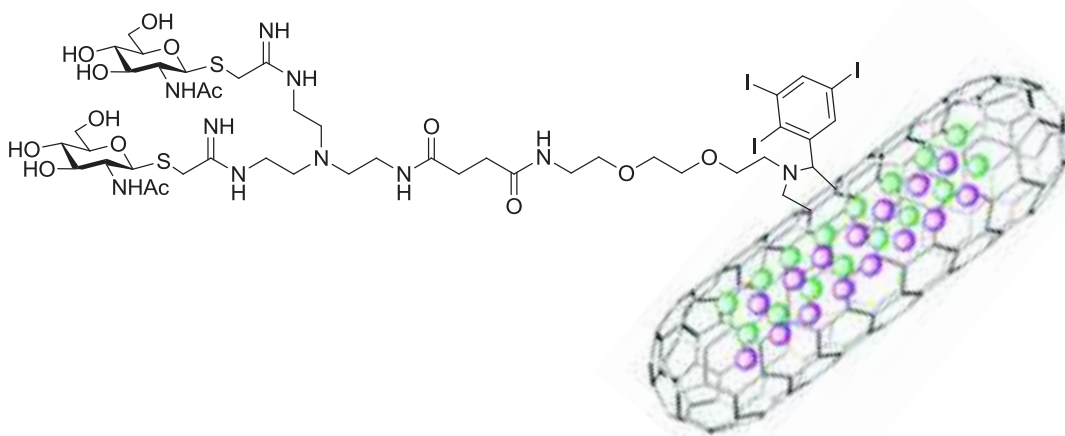
A suspension of **5.44** (9 MBq) in DMF (dry, 3 mL) was sonicated for 2 min, before a solution of **5.18** (12.7 mg, 0.02 mmol, 1 eq.) and **5.4** (9.6 mg, 0.02 mmol, 1 eq.) in DMF (dry, 1 mL) was added. The reaction was refluxed at 130 °C for 4 days, then cooled to r.t. and filtered. The solid was washed with DMF, MeOH and dried under vacuum to afford the title compound **5.45** as a black solid. (RCY = quant.)

(NH₂)₂-linker 2 Na¹²⁵I@SWNTs 5.46



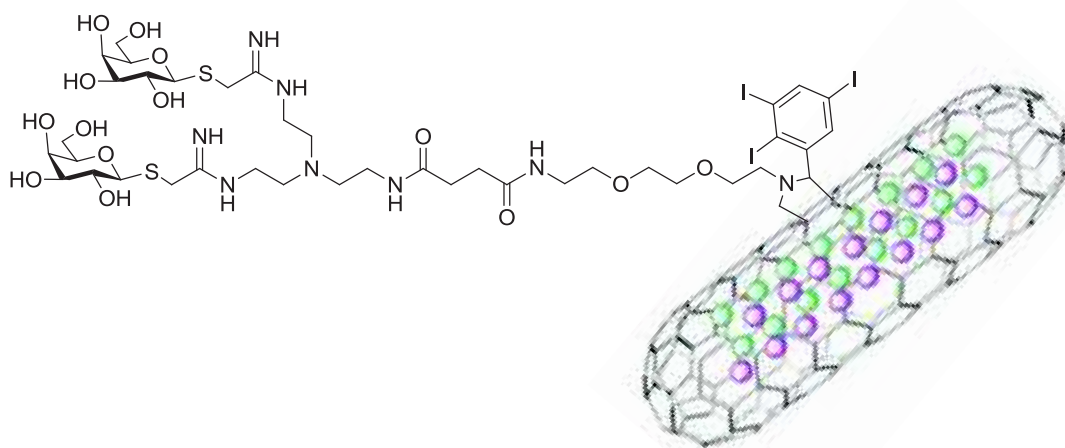
A suspension of **5.45** (9 MBq) in DCM (dry, 5 mL) was sonicated for 2 min, before TFA (2.5 mL) was added. The reaction was stirred at r.t. for 24 h, then evaporated, recovered with MeOH, filtered and washed with MeOH. The solid obtained was dried under vacuum to afford the title compound **5.46** as a black solid. (RCY = quant.)

GlcNAc₂-linker 2 Na¹²⁵I@SWNTs 5.47



A solution of **2.80** (6.5 mg, 20 μ mol, 1 eq.) in MeOH:DMSO (2:1, 1.5 mL) was added to a suspension of **5.46** (3 MBq) in DCM (0.5 mL). The reaction was stirred at r.t. for 30 min. The reaction mixture was centrifuged, the supernatant discarded and the solid recovered with MeOH, filtered and washed with MeOH. The solid obtained was dried under vacuum to afford the title compound **5.47** as a black solid. (RCY = quant.)

Gal₂-linker 2 Na¹²⁵I@SWNTs **5.48**



A solution of **2.77** (6.0 mg, 20 μ mol, 1 eq.) in MeOH:DMSO (2:1, 1.5 mL) was added to a suspension of **5.46** (3 MBq) in DCM (0.5 mL). The reaction was stirred at r.t. for 30 min. The reaction mixture was centrifuged, the supernatant discarded and the solid recovered with MeOH, filtered and washed with MeOH. The solid obtained was dried under vacuum to afford the title compound **5.48** as a black solid. (RCY = quant.)

References

- [1] P. a Boriack, D. W. Christianson, J. Kingery-Wood, and G. M. Whitesides, *J. Med. Chem.* **1995**, 38, 2286-91.
- [2] K. Kordatos, T. Da Ros, S. Bosi, E. Vázquez, M. Bergamin, C. Cusan, F. Pellarini, V. Tomberli, B. Baiti, D. Pantarotto, V. Georgakilas, G. Spalluto, and M. Prato, *J. Org. Chem.* **2001**, 66, 4915-20.
- [3] *Bull. Soc. Chim. Fran.* **1951**, 340.
- [4] R. R. Slama, J., Rando, *Carbohydr. Res.* **1981**, 88, 213-21.
- [5] S. Botros, A. W. Lipkowski, and A. E. Takemori, *J. Med. Chem.* **1986**, 874-76.
- [6] *J. Org. Chem. USSR* **1971**, 7, 2250-53 (English translation).

Biological experiments

6.1 Introduction

6.1.1 Carbohydrate-binding structures

All cell surfaces are covered by a layer of glycoconjugates, which play an important role in physiological and pathological processes through cell-cell communication, inflammation, pathogen recognition and infection.^[1,2] As carriers of information, carbohydrates have far greater potential than the other two major classes of biomolecules, proteins and nucleic acids. Three different nucleotides or amino acids can generate only six unique structures while the same number of monosaccharides can build over a thousand. This level of complexity is one of the reasons why glycans cover so many different roles in biology, and at the same time why their exploration is lagging so far behind. Significant progress has been made thanks to new synthetic strategies^[3] and high-throughput methods,^[4] which allows the synthesis of saccharides of high quality in large quantity, together with the rapid screening of their molecular interactions.

Amongst all existing carbohydrate-binding structures, a specific class of proteins called lectins^[5] (from the Latin 'legere', 'to select') is the most studied. These proteins have the ability to recognise either endogenous carbohydrates or those presented by microbial invaders. Unsurprisingly, lectins are involved in most

carbohydrate-mediated processes.^[6,7] A large number of animal lectins have been isolated and in many cases their roles have been identified. The sugar-binding activity is limited to a portion of most lectins, usually a globular carbohydrate recognition domain (CRD) of fewer than 200 amino acids. These lectins have been grouped into classes based on different features: the nature of their carbohydrate ligands, the biological processes in which they are involved, their subcellular localisation, and their dependence on divalent cations. C-type lectins, named after their requirement of Ca^{2+} ions for their activity, are extracellular and bind a number of sugars through their CRDs. Selectins are a family of cell adhesion molecules similar in many aspects (such as Ca^{2+} dependent binding) to C-type lectins. They all bind to sugars containing both Neu and Fuc such as sLe^x.^[8] S-type (or soluble) lectins are both localised intra- and extracellularly; they do not depend on divalent cations for their activity but often require the assistance of reducing agents (thiols). Their primary ligands are β -galactosides, reason why they are also called galectins.

C-type lectins are found in serum, extracellular matrix and membranes, and are involved in the inflammatory response. There are at least five different varieties of C-type lectins. The first is the type II endocytic receptors, membrane-bound receptors that mediate the endocytosis of glycoproteins. Amongst them are: the asialoglycoprotein receptor (ASGP-R), found on hepatocytes and specific for Gal and GalNAc;^[9] the Kupffer cell receptor, present on the resident macrophages of liver, which binds Fuc and Gal;^[10] the placental mannose receptor, recognising Man and Fuc residues.^[11] The second kind of C-type lectin is the macrophage mannose receptor. This structure is a type I transmembrane protein found on macrophages and hepatic endothelial cells, and mediates the binding and

336

internalisation of glycoconjugates terminating in Man, Fuc and GlcNAc.^[12] Collectins differ from other C-type lectins as they are soluble peptides, consisting of 9-27 units of amino acid units, with COOH-terminating CRDs. They bind sugars on pathogens and play a role in the immune response. Mannose binding proteins (MBPs) A and C are part of this family of receptors. MBP-A binds Fuc and GlcNAc while MBP-C binds the trimannosyl core of N-linked oligosaccharides.^[13] Other collectins are the pulmonary surfactant apoproteins SP-A and SP-D. The first interacts with Man, Glc, Fuc and Gal^[14,15] while SP-D binds only Glc.^[16] Finally, proteoglycans are also C-type lectins. They are found in aggrecan (the cartilage matrix), where they bind Fuc and Gal,^[17] versican (a matrix synthesised by fibroblasts)^[18] and neurocan (found in the brain).^[19] Included in the selectins are E-selectin (in endothelial cells), L-selectin (in leukocytes) and P-selectin (in platelets and endothelial cells). The P-type binds preferentially to Man-6-phosphate. E- and P-selectins on the endothelium interact with neutrophils and leukocytes.^[8] P-selectins on platelets mediate their binding to the endothelium. Galectins play a major role in host defence. Examples of S-type lectins are L-14, L-30 and L-36. L-14 is a homodimer found extracellularly and in cytoplasm. L-30 is a monomer that can be found intracellularly (has a role in RNA processing),^[20,21] extracellularly^[22] and as macrophage surface antigen^[23] binding GlcNAc. L-36 is a polypeptide containing 2 CRDs in tandem.

Binding of carbohydrates occurs in the shallow clefts of the lectins under conditions of rapid flowing blood. Interactions have fast 'on' and 'off' rates of their ligand associations but they are not very strong^[24] (their K_d is in the millimolar range). The structural basis for selective sugar recognition is established by intramolecular forces. Selectivity is achieved through hydrogen bonding with the

sugar hydroxyl groups, van der Waals packing (obtained by the hydrophobic face of the sugar against aromatic side chains of the amino acids in the binding site) and direct and water-mediated contacts between the glycan and the protein surface.

In hydrogen bonding the hydroxyl groups of the sugars act both as donors and acceptors: the sp^3 hybridisation leaves two lone pairs, acceptors of two hydrogen bonds, and a proton, donor of one hydrogen bond. The sugar ring oxygen atom is also sp^3 and has two lone pairs that can act as an acceptor. N-acylation is a recognition determinant of many saccharides as unlike the hydroxyl groups it is not tetrahedral but planar. The amide acts as a hydrogen bond donor while the acetamido oxygen acts as an acceptor. Also, the hydrogen bond can be geometrically specific to epimers (Gal has one axial hydroxyl compared to Glc). Some water molecules present in the binding site of the protein remain after the binding event occurs, forming hydrogen bonds between ligand and receptor. These molecules act as fixed structural elements and therefore can be considered part of the binding site architecture. Recognition of charged groups present on the sugar (e.g. for Neu) often play a limited role in the recognition events, although the formation of salt bridges is a common feature of neuraminidases.^[25-27] Divalent cations (Ca^{2+} , Mn^{2+}) have direct and indirect roles in interactions. They have a direct role when C-type lectins require Ca^{2+} to coordinate sugar ligands through 2 vicinal hydroxyl groups of the sugar: one lone pair of electrons from each hydroxyl moiety coordinates the Ca^{2+} while the other lone pair accepts a hydrogen bond from the amide side chain of the protein.^[28-30] Ca^{2+} and Mn^{2+} have an indirect role when they only fix the position of the sugar

and interacting amino acids in the binding site but do not participate in the binding event themselves.

Sugars are highly polar and solvated but conserve significant nonpolar patches that pack against aromatic side chains in so called stacking interactions. One example is Gal: thanks to the 4-OH in axial position it has an extended non polar surface and as a result it has always been found packed against aromatic residues.

Some lectins show no affinity for monosaccharides but only to complex oligosaccharides. This is the case for selectins and galectins (Gal-binding proteins). Galectins have marginal affinity for Gal alone but bind Lac and LacNAc. This effect can be explained through the hydrogen bonds forming with 3-OH of Glc (or GlcNAc).^[31-33] Tighter binding is observed for LacNAc over Lac because of van der Waals interactions with acetamido group.^[31]

Even when sugars bind in a multivalent manner, there is no cooperativity effect observed. In fact proteins do not change substantially upon binding, nor is there proof of distortions of the sugar ring, while for hydrolytic enzymes the distortion of the ring is part of the enzymatic mechanism.

Other than lectins, a few other proteins interact with sugars, such as enzymes, antibodies, membrane transport proteins and bacterial periplasmic binding proteins.

6.1.2 Carbohydrate-binding specificity

Intense efforts have been spent in defining the ligand specificity of each selectin.^[34] Such information is crucial in understanding the molecular mechanisms behind the biological processes they mediate and in rational design of their antagonists (e.g. to prevent selectin-mediated adhesion). Differences in binding come from the amide position at the binding site, packing, and steric exclusion. Some C-type lectins display selectivity for Man-type ligands (Man and GlcNAc), others for Gal-type ligands (Gal and GalNAc). The mammalian hepatic ASGP-R displays preferential binding to GalNAc as compared to Gal. This is the result of direct van der Waals contact between the His²⁵⁶ and GalNAc which stabilises the sugar in the CRD.^[35] Sugars binding on lectins are also responsible for triggering the inflammation cascade: cells bind to selectin receptors on leukocytes by their surface glycoproteins, containing sLe^x. LacNAc derivatives have been designed as inhibitors of sLe^x expression as potential anti-inflammatory drugs.^[36] Galectin 1 (Gal-1), a lactose-binding lectin, is a component of vascular extracellular matrix and is secreted by human vascular smooth muscle cells (SMCs), particularly abundant in skeletal, cardiac and smooth muscle.^[37,38] Many occlusive vascular diseases occur as a result of abnormal accumulation of smooth muscle cells (SMCs), inflammatory cells and extracellular matrix (ECM) proteins, within the intima of blood vessels.^[39] Migration of SMCs is also involved in the formation of intimal lesions in atherosclerosis, integrins and vascular extracellular matrix assembly.^[40] Gal-1 plays a role in the migration and adhesion of human vascular SMCs and is located intracellularly in ECM of skeletal myofibres. Gal-1 forms homodimers that bind to

340

LacNAc units of polylactosaminoglycans;^[41] similarly, Lac binds to Gal-1^[31] and Gal-2.^[33]

In mammals, fucosylated glycans have pivotal roles in many aspects of biological processes such as lymphocyte homing, immune responses, fertilisation, and development.^[42] Moreover, aberrant fucosylation, which results from the deficiency or overexpression of fucosyl transferases, is associated with a variety of human diseases including cystic fibrosis, leukocyte adhesion deficiency type II, inflammation, cancer progression.^[42-44] L-Fuc significance is manifested in a variety of settings, including the ABO blood group, Lewis-related antigens, core fucosylation of N-linked glycans and O-linked fucose modifications. Disruption of core fucosylation of N-linked glycans leads to disastrous outcomes, as shown by neonatal lethality of mice.^[45] Similarly, disruption of O-fucosylation can have devastating effects on development, culminating in embryonic lethality. Nevertheless, the specific targets responsible for lethality are not yet known.^[46] Sialic acid is another sugar highly involved in both physiological and pathological processes. The functional groups COOH, 4-OH, 8-OH, 9-OH and the acetamido group of sialic acid proved to play an important role in the binding of influenza^[47-49] and sialidase.^[26,50,51] HA and sialidases are involved in virus attachment. HA is responsible for the attachment of viral particles to the host cell through a specific receptor–ligand interaction while the sialidase catalyses the hydrolysis of such ligands to allow the virus to escape from the cell during the budding process.^[52,53]

DC22 is a member of Ig lectins (siglecs) binding Neu- α -2,6-Gal and Neu- α -2,6-Gal-4-GlcNAc sequences. Gal-4 binds Fuc- α -1,2 linked to LacNAc, Gal- α -1,3 and GlcNAc- β -1,3 linked to Lac.^[54] DC-SIGN is a member of the C-type lectin family

which recognises fucosylated and mannosylated glycans (e.g. Le^x and Le^a).^[55] Man interacts with DC-SIGN too,^[56,57] as well as the mannose receptor,^[58] Langerin (CD207) on Langerhans cells,^[58-60] and DC-SIGNR (CD209L) on liver and lymph nodes.^[58,61]

Sialyl-Lewis^x antigen plays an essential role in selectin-dependent leukocyte adhesion and lymphocyte homing. sLe^x has also been implicated in mediating human sperm binding to the zona pellucida in fertilization,^[62] and overexpression of sialyl-Lewis^x is associated to a variety of cancers.^[63] Preliminary conformational studies of sLe^x were performed by ¹H NMR in aqueous solution,^[64] crystal structure to E-selectin^[65] and molecular modelling.^[66,67] sLe^x dissociation constant to E-selectin has been reported to be in the millimolar range.^[67,68]

Chronic inflammatory diseases and autoimmune pathologies are the result of the body's overzealous recruitment of lymphocytes from the circulation to a location suspected to be the site of an injury. White cells are dedicated to the repair of tissue damage and defence against microbial infection. Leukocytes are brought to the site of the injury by a series of steps referred to as the inflammatory cascade. Damaged tissues release cytokines that stimulate the endothelium to express E- and P-selectins,^[69,70] which recognise sLe^x on the surface of leukocytes^[71] and promote their adhesion to the affected endothelial cells.^[72-74] E-selectins recognise sLe^x,^[75-78] sLe^a^[75] and their sulphated derivatives,^[79,80] same as per L-selectins.^[81-83] P-selectins recognise both sLe^x and sLe^a but not their sulphated derivatives.^[84,85] The role of Lewis oligosaccharides in inflammation is not just involved in triggering the response cascade: recombinant thrombomodulin domain 1 (rTMD1) binds to lipopolysaccharide (LPS) by interacting with Le^y and neutralises the inflammatory response induced by LPS

342

(leading cause of sepsis).^[86] DC-SIGNs recognise antigens bearing mannose-rich or fucosylated glycans; Rouhanifard et al.^[87] captured DC-SIGN positive cells in a mixed cell population with nanoparticles functionalised with sLe^x, LacNAc, and Lac.

6.2 Glycomap *in vivo* experiments

The glyconanoparticles (GNPs) comprising the Glycomap have been tested *in vivo* to establish whether and how different sugars affect their distribution in the body.

6.2.1 *In vivo* experimental set up

Firstly, we investigated different conditions to set standard parameters in testing the Glycomap. The three main variables we considered were the dose injected, coverage of the GNPs and circulating time before harvesting the organs. In order to quantify the amount of NPs accumulating in each organ we excised a section from each organ, homogenised the tissue, separated the NPs by using a magnet and counted the NPs with a haemocytometer. NPs distribution has been represented as % on relative counting between organs.

Some organs, liver and spleen in particular, have been identified as clearance organs, meaning those organs which represent the preferential route of the body for the excretion of the NPs. They were determined after injecting non-functionalised NPs (NanoLink, Fig. 6.1).

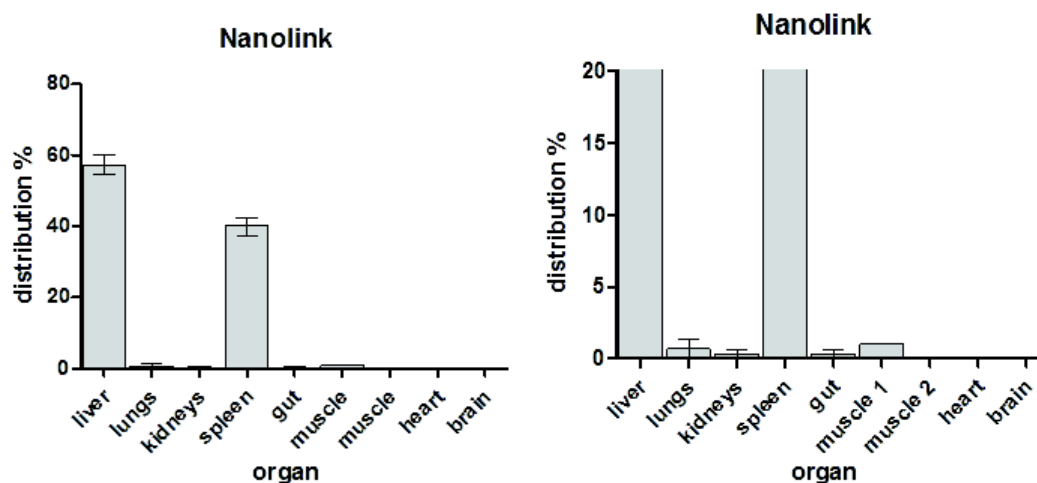


Figure 6.1 NanoLink NPs distribution in CD-1 mice (left) and enhancement (right).

In order to see an enrichment of each GNP in non-clearance organs we needed to inject enough NPs to saturate the liver and spleen. Therefore, we ran a test injecting mice with 4 mg/Kg of GlcNAc-NPs and in parallel, 4 mg/Kg of NanoLink followed by a second injection of 4 mg/Kg of GlcNAc-NPs. NPs were isolated from organs and counted (Fig. 6.2).

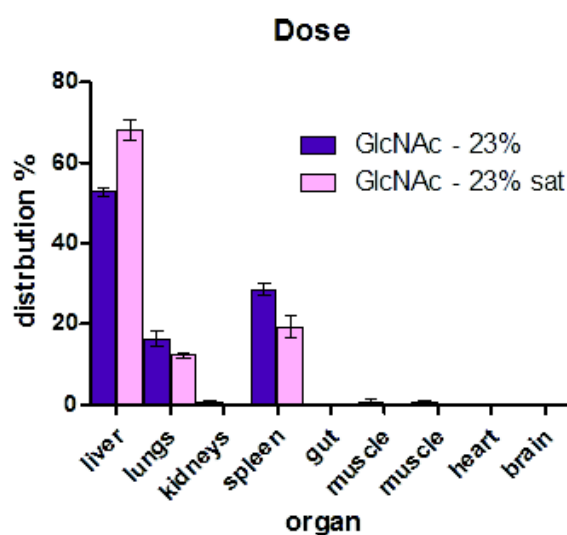


Figure 6.2 GlcNAc-NPs distribution in CD-1 mice alone (purple) and after pre-injection of NanoLink (pink).

As shown in Figure 6.2, the first injection did not affect significantly the distribution of the sugar-coated NPs, suggesting that by using a dose of 4 mg/Kg, we were already operating at saturation conditions.

Mice were injected with Lac-NPs bearing different sugar coverage (14-35-68 %), then isolated and counted (Fig. 6.3).

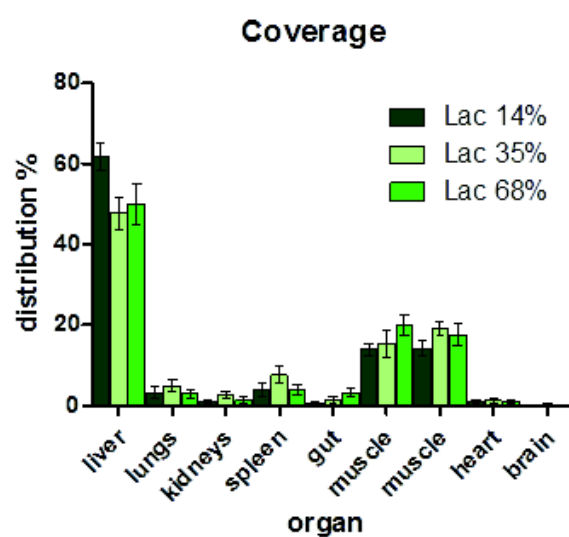


Figure 6.3 Lac-NPs distribution in CD-1 mice with 3 different coverages of GNPs.

From the results obtained the difference in % of coverage does not seem to profoundly affect the particles' distribution, so we tested the GNPs we prepared at 65 ($\pm 10\%$) coverage, to maintain consistency.

Finally, Lac-NPs were injected and left to circulate in the animals for 1 h and 3 h before being isolated and counted (Fig. 6.4).

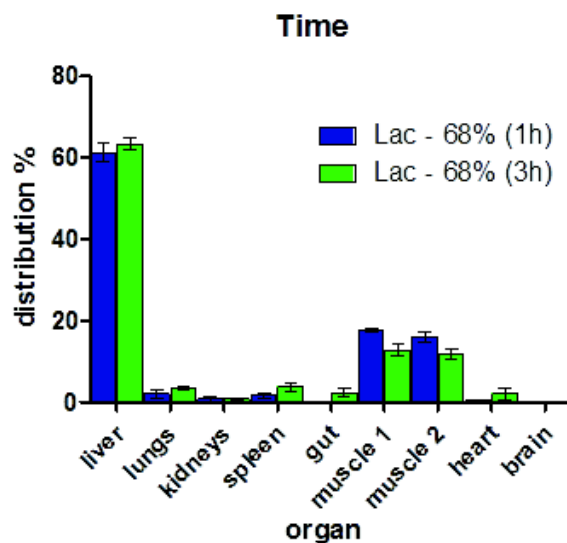


Figure 6.4 Lac-NPs distribution in CD-1 mice when culling the animals after 1 h (blue) and 3 h (green).

From the data obtained not even a longer circulation time affected the distribution significantly.

Based on the experiments described above, it was decided to test all the GNPs of Glycomap with a coverage of 65% ($\pm 10\%$) injecting 4 mg/Kg mouse in 100 μL of saline, and culling the animals 3h after the injection and recovery.

Other parameters were set independently; CD-1 male mice were selected as candidates for our experiments. Thanks to their low body weight, less GNPs were required per animal. Samples were administered intravenously by tail vein injection. Mice were killed by injection of pentobarbital and perfused to remove blood from the organs. We decided to limit the quantification of the distribution of the GNPs to 9 different organs (liver, lungs, heart, kidneys, spleen, gut, muscles from both legs and brain). All experiments were run in triplicate, from the same batch of samples for each compound. Mice were injected with saline solution (0.9%) and with non-modified nanoparticles, as controls.

6.2.2 GNPs distribution and statistical analysis

All the GNPs synthesised for the Glycomap in chapter 4 were tested following the conditions set in the previous section. The data obtained were analysed using one-way and two-way ANOVA (Analysis Of Variance), and by unpaired two-tailed Student's t-test to confirm their significance ($P < 0.05$ was considered statistically significant).

Sugar distribution was compared to non-coated NanoLink for each organ. Most of the samples, as expected, were found mainly in clearance organs, especially in the liver (Fig. 6.5).

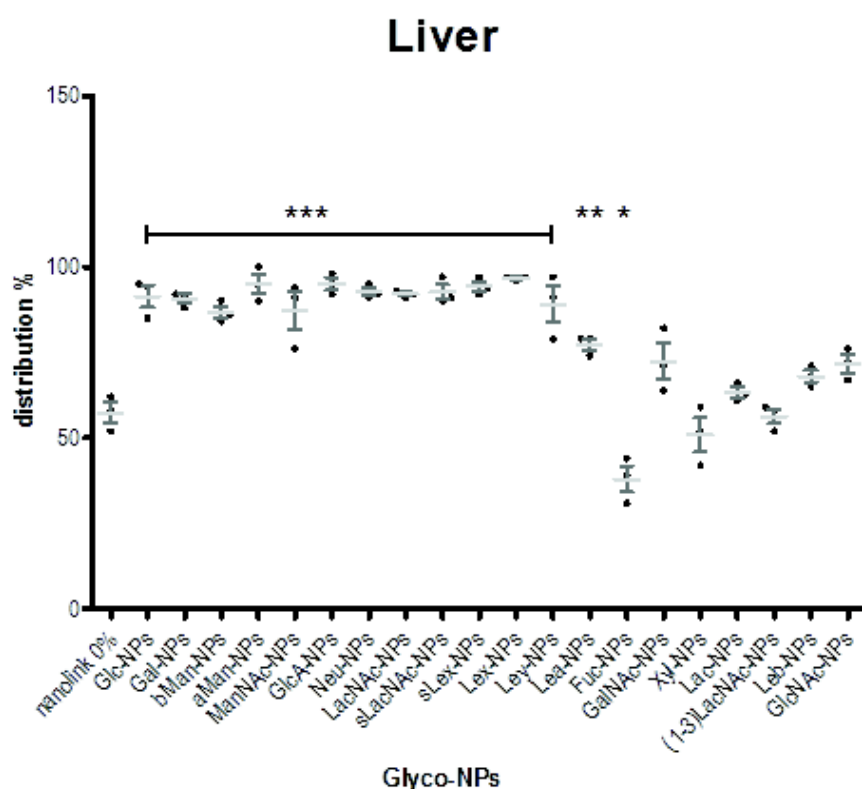


Figure 6.5 Comparison of the GNPs' distribution in the liver of CD-1 mice.

Five sugars, Fuc, GalNAc, Xyl, Le^b and 1,3-LacNAc, had a distribution similar to NanoLink, although with a different ratio between liver and spleen. While Xyl matched NanoLink (liver ~50-60%, spleen ~40%), Fuc showed an inversion of the values (liver ~40%, spleen ~60%). GalNAc, Le^b and 1,3-LacNAc did not deviate significantly either, but showed a slight increased accumulation in the liver with respect to the spleen (liver ~60-70%, spleen ~20-30%). The two-tailed P value for each sugar distribution in the liver differing from NanoLink was calculated by unpaired t-test (Table 6.1).

SAMPLE	P VALUE
Glc	0.0014
Gal	0.0005
βMan	0.001
αMan	0.0008
ManNAc	0.0091
Fuc	0.0155
GlcA	0.0004
Neu	0.0004
1,4-LacNAc	0.0003
sLacNAc	0.0006
sLe ^x	0.0003
Le ^x	0.0002
Le ^y	0.0063
Le ^a	0.0063

Table 6.1 P value of GNPs distribution in the liver respect to NanoLink.

As a consequence of the major accumulation in the liver, fewer particles were found in the other clearance organ, the spleen, with respect to NanoLink (Fig. 6.6).

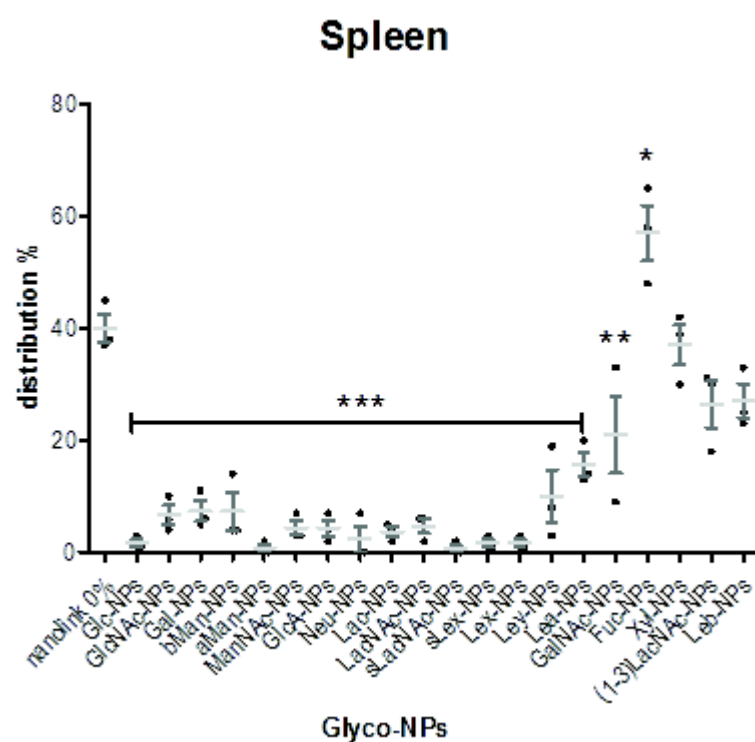


Figure 6.6 Comparison of the GNPs' distribution in the spleen of CD-1 mice.

The two-tailed P value for the sugars' distribution in the spleen was calculated again (Table 6.2).

Sample	P value
Glc	0.0001
Gal	0.0005
βMan	0.0014
αMan	0.0001
GlcNAc	0.0004
ManNAc	0.0002
Fuc	0.0373
GlcA	0.0003
Neu	0.0004
Lac	0.0002
1,4-LacNAc	0.0002
sLacNAc	0.0001
sLe^x	0.0001
Le^x	0.0001
Le^y	0.0005
Le^a	0.0019

Table 6.2 P value of GNPs' distribution in the spleen with respect to NanoLink.

The most interesting results appeared from the analysis of the enrichment in non-clearance organs, where a minority of the GNPs accumulated. Here we found that some sugars presented a significantly different accumulation in some organs. GlcNAc accumulation in the lungs (Fig. 6.7) was higher than for the NanoLink (17% against 1% for NanoLink, P value = 0.0011).

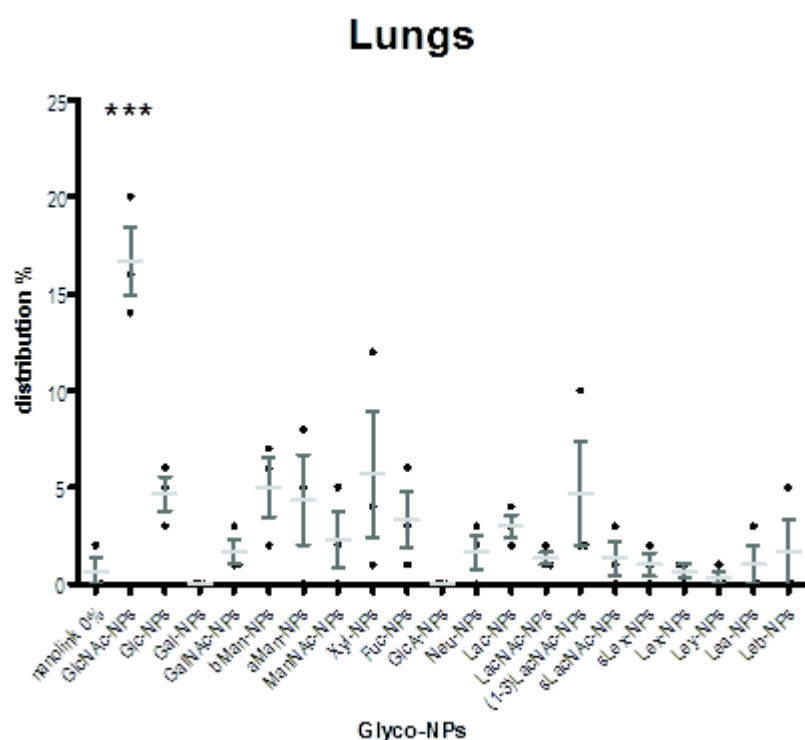


Figure 6.7 Comparison of the GNPs' distribution in the lungs of CD-1 mice.

An unexpected accumulation of Lac-NPs in the front leg muscle (muscle 2) of mice was consistently observed (Fig. 6.8). The Lac-NPs counted were 21% of the total NPs injected whereas no particles were found in the muscle for NanoLink (P value = 0.0222).

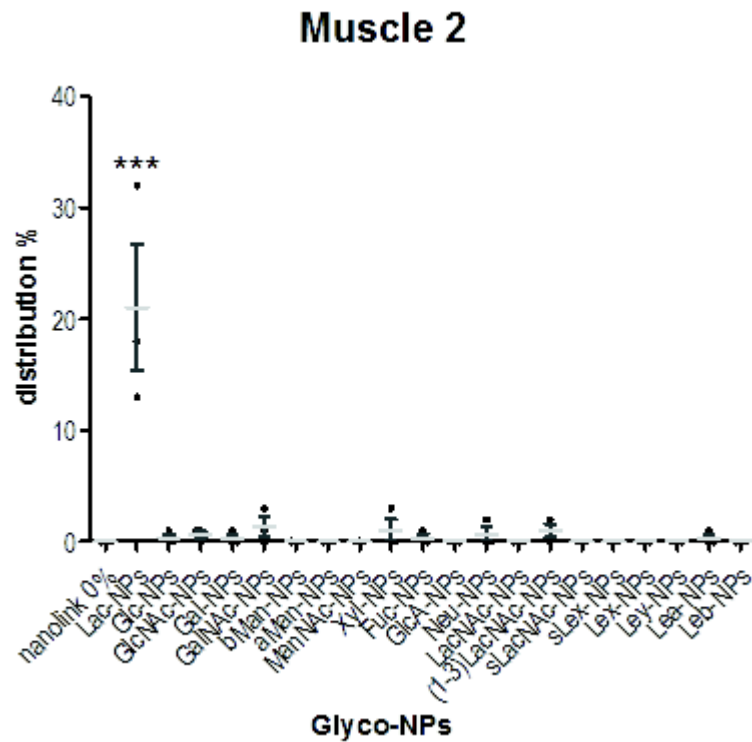


Figure 6.8 Comparison of the GNPs' distribution in the front leg muscle of CD-1 mice.

To confirm these surprising results we collected samples from a different muscle, to be precise from the hind leg (mucle 1, Fig. 6.9). On this occasion the accumulation of Lac-NPs compared to NanoLink still persisted (12% against 1%, P value = 0.0043) but was accompanied by 1,3-LacNAc-NPs (10% against 1% of NanoLink, P value = 0.0015).

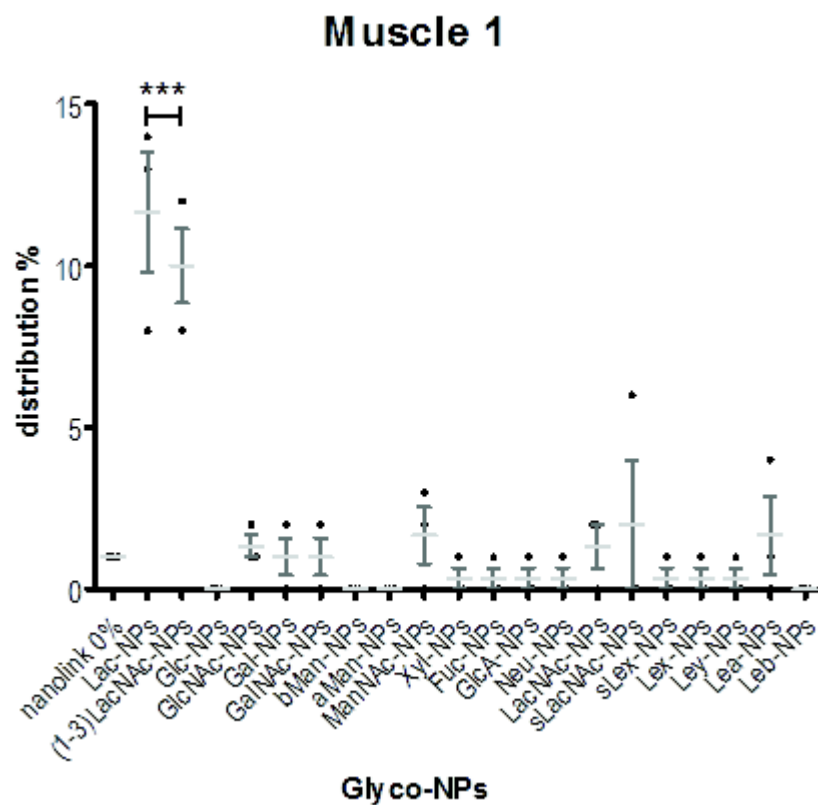


Figure 6.9 Comparison of the GNPs' distribution in the hind leg muscle of CD-1 mice.

One last set of analyses was run to evaluate the effect of the chemical variance between sugar structures on their *in vivo* distribution. The first variable considered was the anomericity of α and β Man (Fig. 6.10).

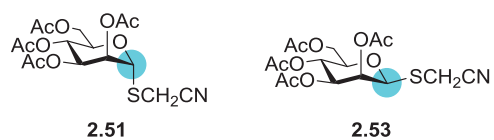


Figure 6.10 Structures of α (left) and β (right) Man.

Unfortunately no significant difference between their distributions was observed. We proceeded by looking at the difference between Glc and its oxidised analogue GlcA (Fig. 6.11).

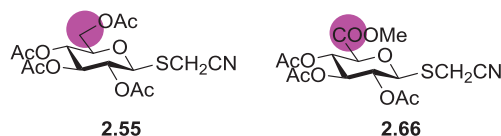


Figure 6.11 Structures of Glc (left) and GlcA (right).

Again no difference between each Glc analogue's distribution was observed. We finally came across a variation where we looked at Glc and its deoxy analogue Xyl (Fig. 6.12).

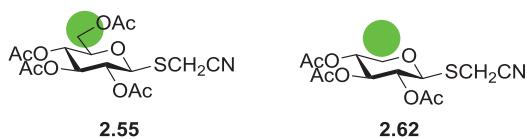


Figure 6.12 Structures of Glc (left) and Xyl (right).

Even though a difference between the distribution of the two structures was observed, this was limited to the liver and spleen only (Fig. 6.13, P value = 0.0023 and 0.0006 respectively).

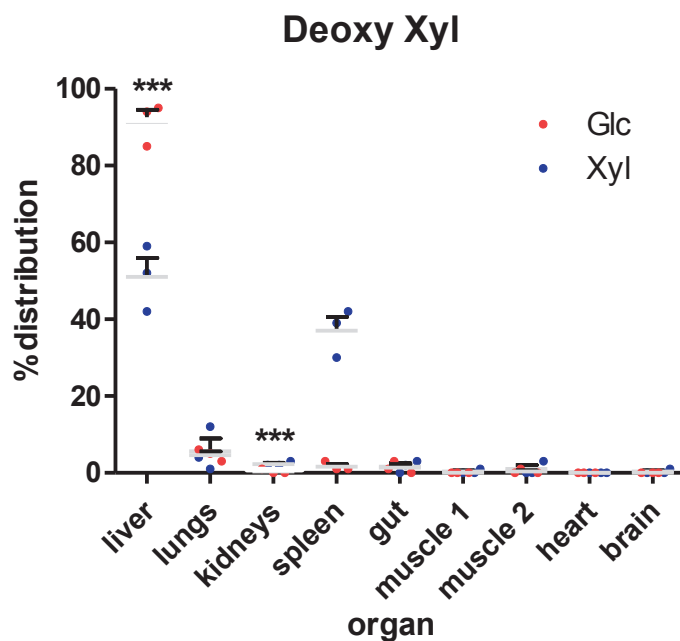


Figure 6.13 Comparison of distribution of Glc and Xyl in CD-1 mice.

At this point, we directly compared all the epimers in our Glycomap: Gal with Glc, β -Man with Glc, and GalNAc with GlcNAc (Fig. 6.14).

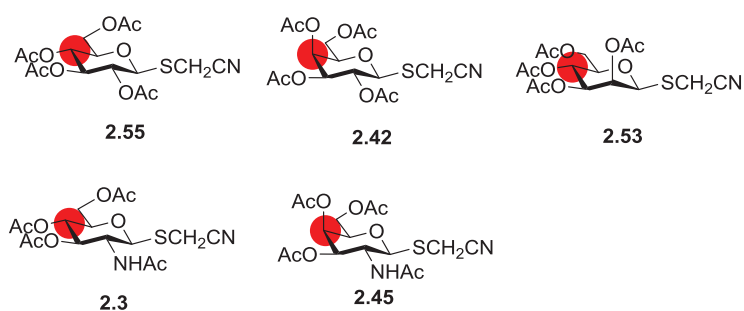


Figure 6.14 Structures of Glc, Gal, Man (top, from left), GlcNAc and GalNAc (below, from left).

We could still not detect any variation amongst them, so we moved on to another chemical modification very common in nature, N-acetylation. Here we compared

GalNAc with Gal, GlcNAc with Glc, α -Man with ManNAc, and 1,4-LacNAc with Lac (Fig. 6.15).

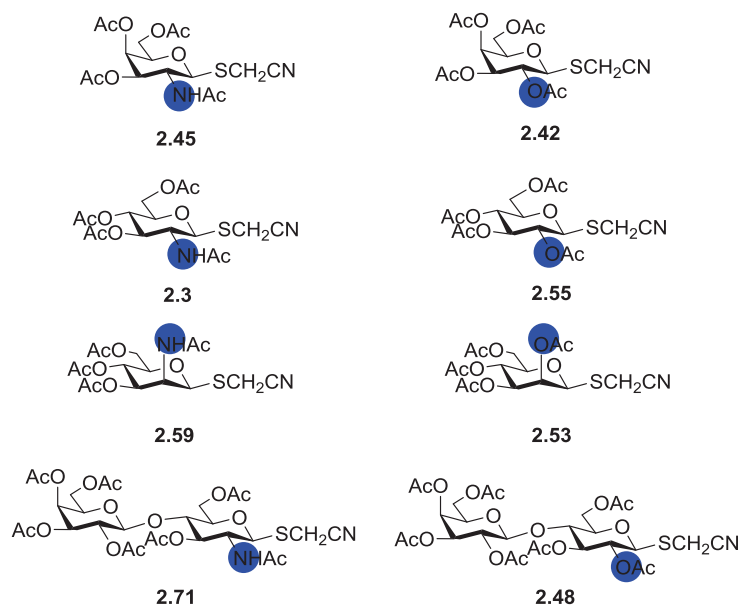


Figure 6.15 Structures of GalNAc **2.45**, Gal **2.42**, GlcNAc **2.3**, Glc **2.55**, α -Man **2.59**, ManNAc **2.53**, 1,4-LacNAc **2.71** and Lac **2.48**.

N-acetylation had an effect in the distribution of almost all the analogues involved (Fig. 6.16).

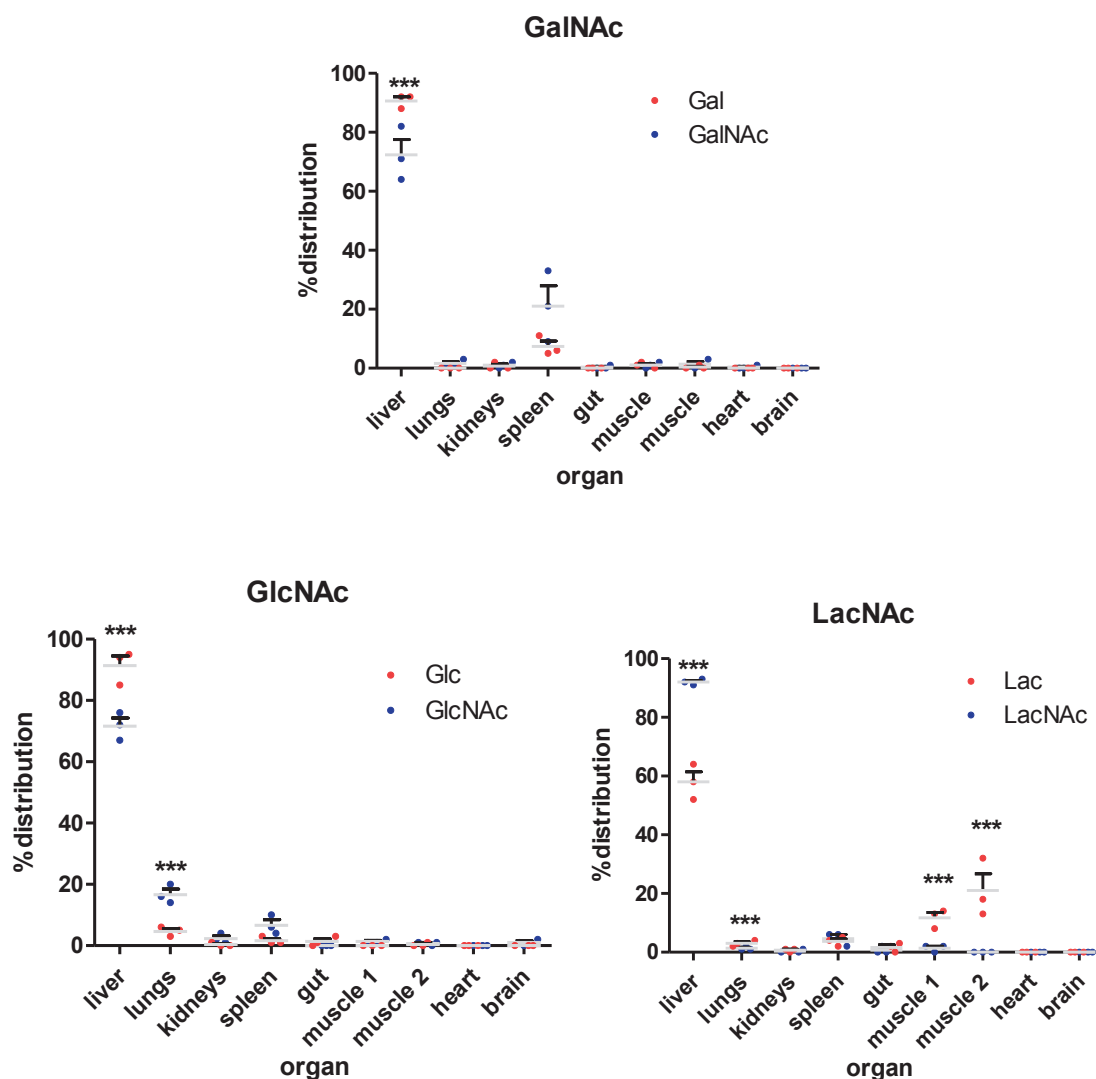


Figure 6.16 Comparison of distribution of GalNAc, Gal, GlcNAc, Glc, 1,4-LacNAc and Lac in CD-1 mice.

Gal and GalNAc had different distributions in the liver (P value = 0.0275), while Glc and GlcNAc had different distributions in both the liver and the lungs (P value = 0.0087 and 0.0037 respectively). Finally Lac and 1,4-LacNAc differed in liver, muscle 1 and muscle 2 (P value = <0.0001, 0.0063 and 0.0222 respectively). The last two variables we compared were the introduction of Fuc and Neu. Fucosylation was considered for 1,4-LacNAc over Le^x , 1,3-LacNAc over Le^a , Le^x over Le^y , Le^a over Le^b and sLacNAc over s Le^x (Fig.6.17).

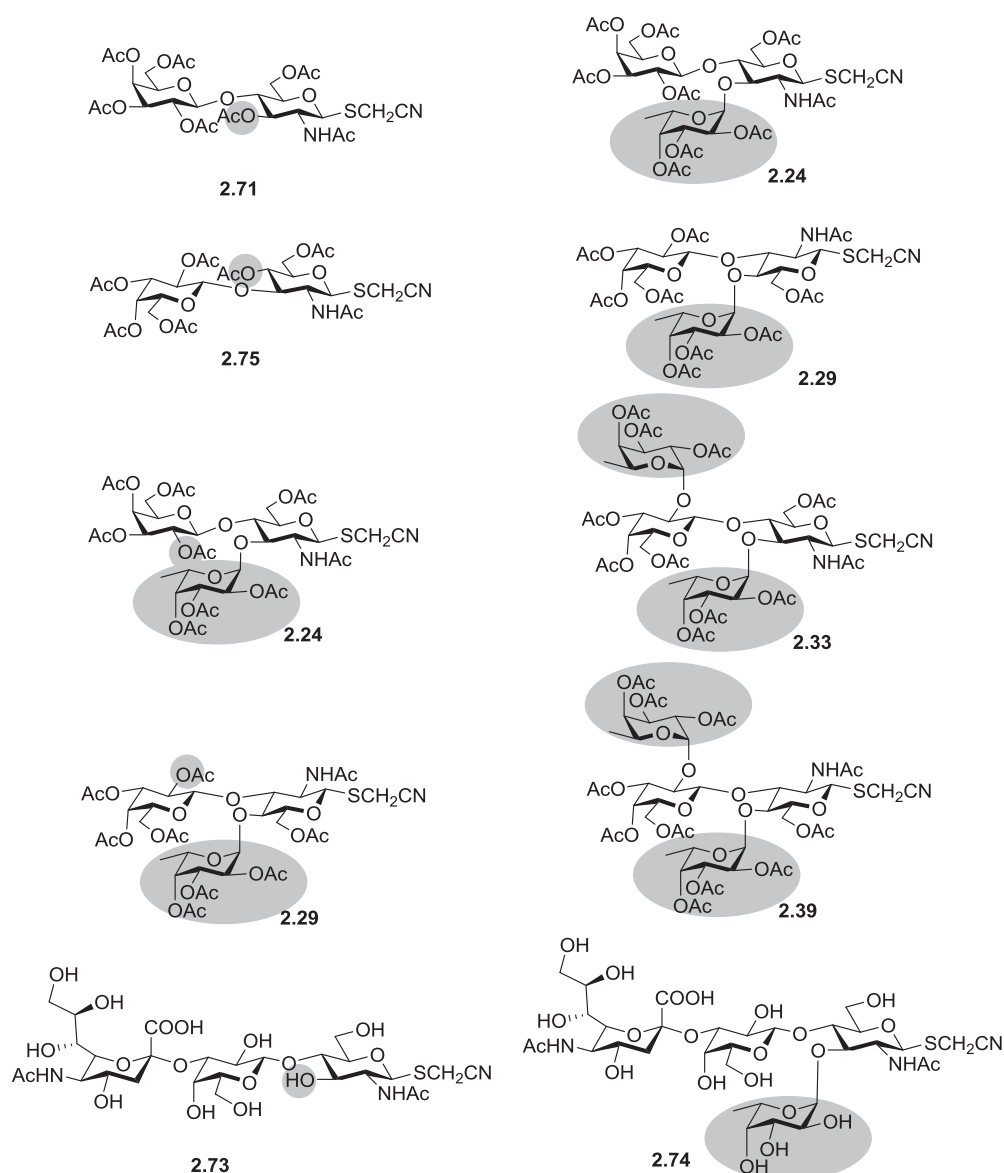


Figure 6.17 Structures of 1,4-LacNAc **2.71**, Le^x **2.24**, 1,3-LacNAc **2.75**, Le^a **2.29**, Le^y **2.33**, Le^b **2.39**, sLacNAc **2.73** and sLe^x **2.74**.

Fucosylation seemed to have an effect only on Le^a over 1,3-LacNAc in liver and muscle 1 (Fig. 6.18, P value = 0.0016 and 0.0007 respectively).

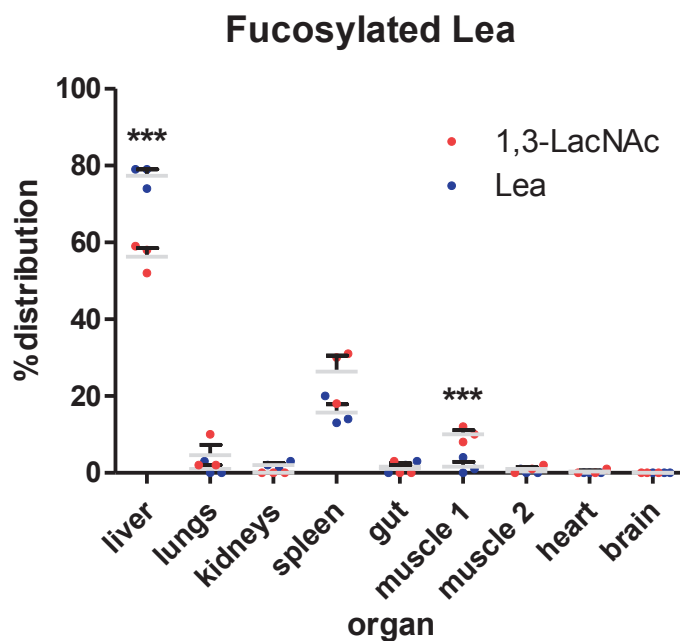


Figure 6.18 Comparison of distribution of 1,3-LacNAc and Le^a in CD-1 mice.

Sialylation of 1,4-LacNAc to sLacNAc and of Le^x to sLe^x (Fig. 6.19) had no effect on the tetrasaccharides' distribution.

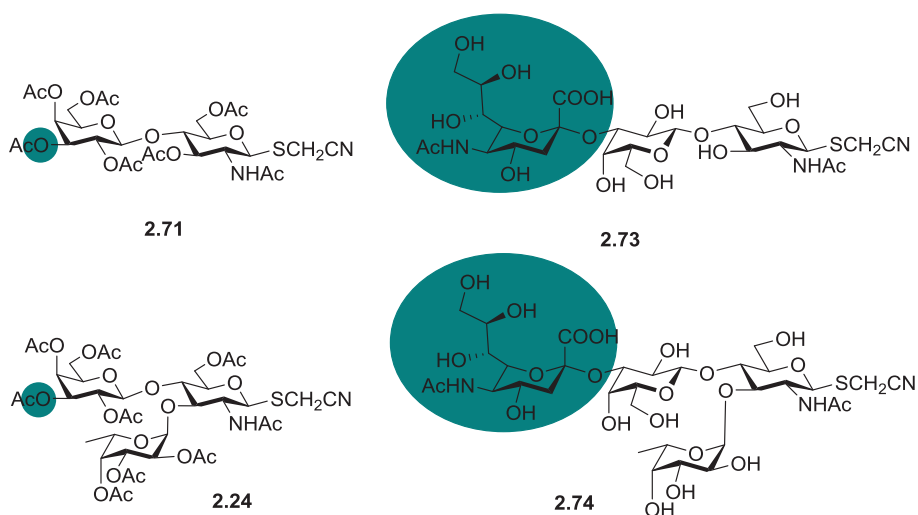
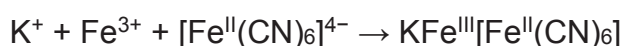


Figure 6.19 Structures of 1,4-LacNAc **2.71**, sLacNAc **2.73**, Le^x **2.24**, and sLe^x **2.74**.

6.2.3 Histological studies

To confirm the results we obtained with the isolation and counting of the NPs, we looked into a technique that would have allowed us to collect images of the NPs accumulated in the tissues. Sections of organs excised post-mortem were embedded in wax and cut in slices for histochemical staining. Slices of tissue were mounted on slides and stained by Perl's staining, followed by nuclear fast red. Perl's staining (or Prussian blue) is based on the reaction of potassium ferrocyanide with the iron (Fe^{3+}) of the NPs:



This reaction generates ferricyanide, a blue pigment that allows the detection of the iron NPs. Samples were counterstained with nuclear fast red to provide a good contrast and identify the organs through their characteristic structural features. Perl's staining located the NPs accumulated in the organs that were previously observed by counting (Fig. 6.20)

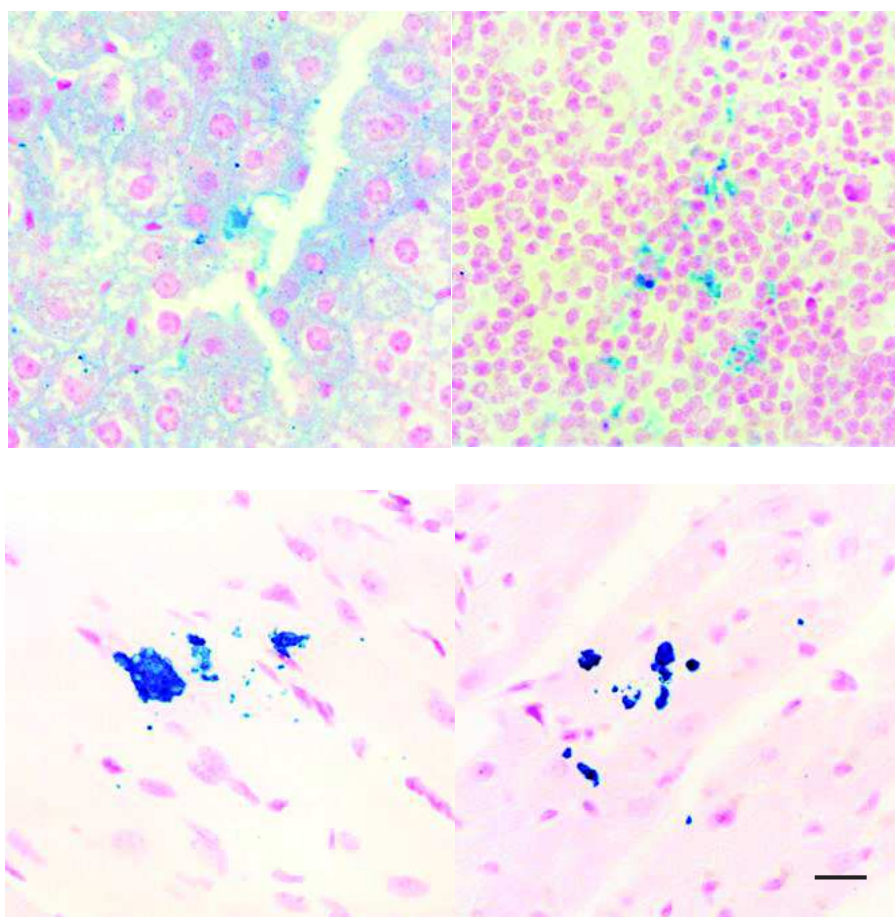


Figure 6.20 Histological staining of liver (top left), spleen (top right), muscle 2 (bottom left) and muscle 1 (bottom right) containing NPs. Scale bar is 100 μm .

6.3 *In vivo* testing of ^{18}F -GNPs

The ability of Le^{x} -NPs to bind to inflamed endothelium was already assessed by Sander et al.^[88] IL-1 β was selected to provoke inflammation, with upregulation of P- and E- selectins, because of its capability of activate the endothelium without causing a breakdown of the blood-brain barrier,^[89-91] preventing an aspecific accumulation of the contrasts. Thus, IL-1 β was directly injected into the left striatum of the brain of rats and left for 4 h to develop a localised inflammation.^[90] The contralateral cranial hemisphere remained untreated to provide an internal

control. GNPs labelled with ^{18}F and decorated with Le^x (**4.24**) and Le^a (**4.25**) were injected *via* tail-vein after inflammation developed. Animals were culled after 1 h, the brains frozen instantly to be cut in slices with a cryostat. The slices, deposited on a phosphorous screen, were subjected to autoradiography (Fig. 6.21).

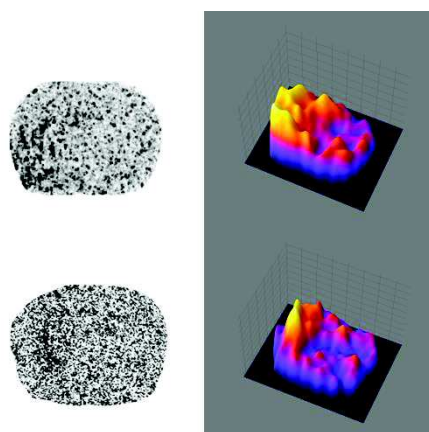


Figure 6.21 Autoradiography of brain sections after injection with $[^{18}\text{F}]\text{-Le}^a\text{-NPs}$ **4.25** (top) and $[^{18}\text{F}]\text{-Le}^x\text{-NPs}$ **4.24** (bottom).

As per the preliminary results obtained before (see chapter 4), Le^a NPs accumulated in the site of inflammation, while Le^x accumulation was marginal. No NPs were detected in the right hemisphere, where no inflammation was provoked.

6.4 *In vivo* testing of glycosylated $\text{Na}^{125}\text{I}@\text{SWNTs}$

$\text{Na}^{125}\text{I}@\text{SWNTs}$ glycosylated with GlcNAc (**5.47**) and Gal (**5.48**) were tested together with $\text{Na}^{125}\text{I}@\text{SWNTs}$ functionalised with linker only (**5.46**). The organs (liver, spleen, brain, heart, muscle 1, muscle 2, gut, kidneys and lungs) were collected and their activity measured by γ counter, following the example of what

was previously reported for ^{125}I -labelled NPs.^[92] The activity was quantified as percentage of injected dose per gram of animal (%ID/g). We acknowledge that we could not measure the activity of the remaining carcasses, therefore part of the activity might have been lost with them and our calculations could be affected by a significant error. However, after the experiments the activity detected on the carcasses was negligible, therefore we consider the results obtained still significant. As previously reported^[93] GlcNAc-targeted hot NTs accumulated in the lungs, without sign of radioactivity in any other organ. Unfortunately, the same distribution was obtained for non-glycosylated and galactosylated NTs as well (Table 6.2).

organ	5.46	5.47	5.48
liver	2	2	2
lungs	85	84	86
kidneys	2	2	2
spleen	3	2	2
gut	2	2	2
muscle 1	2	2	2
muscle 2	1	2	2
heart	1	2	2
brain	1	1	1

Table 6.2 Activity measured by γ counter of organs harvested from CD-1 mice after injection of **5.46**, **5.47** and **5.48**.

These results, give a clear indication that this system is not ready to be used in testing the sugars of the Glycomap, and even less to deliver radioactivity. The value of these data was vital as prevented us from testing other sugars until the system is deemed suitable for the purpose.

6.5 Conclusions

Within this chapter we concluded the exploration of the systems we designed, prepared and described in the previous chapters, by testing them *in vivo* (with the exception of the glycosylated proteins).

GNPs prepared to build the Glycomap have been injected into mice and their distribution quantified by isolating and counting the NPs in several organs. Most of the GNPs accumulated in the liver and, to some extent, in the spleen, organs identified to be the preferential routes for the excretion of NPs. Besides the uptake in the liver and spleen, some targeting was still achieved. Firstly the partial accumulation of GlcNAc-modified NPs in the lungs, which is not entirely unexpected as GlcNAc functionalised nanoparticles (even though of different composition) have been seen to target lungs before.^[93] Secondly, and more surprisingly, the accumulation of Lac- and 1,3-LacNAc-NPs in the striated muscle: sections of muscle were collected from two different sites (front and hind legs), and in both NPs were detected. There is a limited evidence for muscle targeting ^[94-96] and therefore these findings will need further investigation to understand the dynamics of this interaction and possibly isolate the interacting partner. All results have been confirmed by histological studies of the sections collected. We suggest that these GNPs can be used to explore further conditions, interactions and ligands. We do not claim this to be a universal system, considering the effect that all the features of the NPs (size, charge, coating, etc.), the linker and the limitations we had to set for our experiments, have on the results we obtained. We believe however that this system represents a solid

platform to start our exploration of the glycode, and hope that future research will be inspired from it and diversify in other systems.

The testing of the radiolabelled-glyconanotubes did not show results as favourable. GlcNAc-NTs targeted the lungs as reported before^[93] but the non-glycosylated and Gal-NTs, disappointingly, shared the same fate. Our hypothesis that Gal would have targeted them in the liver proved inconsistent and so the fact that GlcNAc is responsible for the targeting in the lungs is questionable. The data suggests that the targeting observed before might have been erroneously assigned to the GlcNAc while the real culprit could be the nanotubes themselves, more due to their size rather than from any active targeting. We believe that nanotubes still hold great potential but necessary optimisations regarding mainly their dispersibility have to take place before they can be used for any imaging or therapeutic purposes.

One last topic that has not been addressed during the biological studies of these systems is the fate that they will share after the accumulation in the organs. Although some processes promoting the elimination of NPs and CNTs have been described^[97-99] we did not pursue any investigation of this sort, and we recognise this to be a very critical point in the development of nanosystems for clinical use.

6.6 References

- [1] R. A. Dwek, *Chem. Rev.* **1996**, 96, 683-720.
- [2] A. Varki, *Glycobiology* **1993**, 3, 97-130.
- [3] P. H. Seeberger and D. B. Werz, *Nature* **2007**, 446, 1046-51.
- [4] S. Fukui, T. Feizi, C. Galustian, A. M. Lawson, and W. Chai, *Nat. Biotechnol.* **2002**, 20, 1011-7.
- [5] N. Sharon, *J. Biol. Chem.* **2007**, 282, 2753-64.
- [6] H. Ghazarian, B. Idoni, S. B. Oppenheimer, *Acta Histochem.* **2011**, 113, 236-47.
- [7] N. Sharon, H. Lis, *Essays Biochem.* **1995**, 30, 59-75.
- [8] L. A. Lasky, *Science* **1992**, 258, 964-69.
- [9] M. Spiess, *Biochemistry* **1990**, 29, 10009-18.
- [10] G. W. Hoyle and R. L. Hill, *J. Biol. Chem.* **1988**, 263, 7487-92.
- [11] B. M. Curtis, S. Scharnowske, and A. J. Watson, *Proc. Natl. Acad. Sci. U. S. A.* **1992**, 89, 8356-60.
- [12] P. D. Stahl, *Am. J. Respir. Cell. Mol. Biol.* **1990**, 2, 317-8.
- [13] R. A. Childs, T. Feizi, C. T. Yuen, K. Drickamer, and M. S. Quesenberry, *J. Biol. Chem.* **1990**, 265, 20770-7.
- [14] H. P. Haagsmansq, S. Hawgood, T. Sargeants, R. T. Whitell, K. Drickamer, and B. J. Benson, *J. Biol. Chem.* **1987**, 262, 13877-80.
- [15] R. A. Childs, J. R. Wright, G. F. Ross, C-T. Yuen, A. M. Lawson, W. Chai, K. Drickamer, and T. Feizi, *J. Biol. Chem.* **1992**, 267, 9972-9.
- [16] A. Persson, D. Chang, and E. Crouch, *J. Biol. Chem.* **1990**, 265, 5755-60.

- [17] K. J. Doege, M. Sasaki, T. Kimura, and Y. Yamada, *J. Biol. Chem.* **1991**, 266, 894-902.
- [18] D. R. Zimmermann and E. Ruoslahti, *EMBO J.* **1989**, 8, 2975-81.
- [19] U. Rauch, L. Karthikeyan, P. Maurel, R. U. Margolis, and R. K. Margolis, *J. Biol. Chem.* **1992**, 267, 19536-47.
- [20] S. Jia and J. L. Wang, *J Biol Chem* **1988**, 263, 6009-11.
- [21] J. L. Wang, J. G. Laing, and R. L. Anderson, *Glycobiology* **1991**, 1, 243-52.
- [22] A. Raz, G. Pazerini, and P. Carmi, *Cancer Res.* **1989**, 49, 3489-93.
- [23] B. J. Cherayil, S. J. Weiner, and S. Pillai, *J. Exp. Med.* **1989**, 170, 1959-72.
- [24] R. Alon, D. A. Hammer, and T. A. Springer, *Nature* **1995**, 374, 539-542.
- [25] W. P. Burmeister, R. W. Ruigrok, and S. Cusack, *EMBO J.* **1992**, 11, 49-56.
- [26] J. N. Vargheese, W. G. Laver, and P. M. Colman, *Nature* **1983**, 303, 35-40.
- [27] S. Crennell, E. Garman, G. Laver, E. Vimr, and G. Taylor, *Structure* **1994**, 2, 535-44.
- [28] W. I. Weis, K. Drickamer, and W. A. Hendrickson, *Nature* **1992**, 360, 127-34.
- [29] A. R. Kolatkar and W. I. Weis, *J. Biol. Chem.* **1996**, 271, 6679-85.
- [30] K. K-S. Ng, K. Drickamer, and W. I. Weis, *J. Biol. Chem.* **1996**, 271, 663-74.
- [31] D. I. Liao, G. Kapadia, H. Ahmed, G. R. Vasta, and O. Herzberg, *Proc. Natl. Acad. Sci. U. S. A.* **1994**, 91, 1428-32.

- [32] Y. Bourne, B. Bolgiano, D-I. Liao, G. Strecker, P. Cantau, O. Herzberg, T. Feizi, C. Cambillau, *Nat. Struct. Mol. Biol.* **1994**, 1, 638-53.
- [33] Y. D. Lobsanov, M. A. Gitt, H. Leffler, S. H. Barondes, and J. M. Rini, *J. Biol. Chem.* **1993**, 268, 27034-8.
- [34] C. Laudanna, G. Constantin, P. Baron, E. Scarpini, G. Scarlato, G. Cabrini, C. Dechecchi, F. Rossi, M. A. Cassatella, and G. Berton, *J. Biol. Chem.* **1994**, 269, 4021-6.
- [35] A. R. Kolatkar, A. K. Leung, R. Isecke, R. Brossmer, K. Drickamer, and W. I. Weis, *J. Biol. Chem.* **1998**, 273, 19502-8.
- [36] T. K-K. Mong, L. V. Lee, J. R. Brown, J. D. Esko, and C.-H. Wong, *Chembiochem* **2003**, 4, 835-40.
- [37] S. H. Barondes et al., *Cell* **1994**, 76, 597-8.
- [38] B. S. S. Rabinowitz and S. Gordon, *J. Exp. Med.* **1991**, 174, 827-36.
- [39] G. H. Gibbons and V. J. Dzau, *Science* **1996**, 272, 689-93.
- [40] S. J. Shattil and M. H. Ginsberg, *J. Clin. Invest.* **1997**, 3, 1-5.
- [41] Q. Zhou and R. D. Cummings, *Arch. Biochem. Biophys.* **1990**, 281, 27-35.
- [42] D. J. Becker and J. B. Lowe, *Glycobiology* **2003**, 13, 41R-53R.
- [43] E. Miyoshi, K. Moriwaki, and T. Nakagawa, *J. Biochem.* **2008**, 143, 725-9.
- [44] C. Chen, Y. Jan, Y. Juan, C-J. Yang, M. Huang, C. Yu, P. Yang, M. Hsiao, T-L. Hsu, and C-H. Wong, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, 110, 630-5.
- [45] X. Wang, S. Inoue, J. Gu, E. Miyoshi, K. Noda, W. Li, Y. Mizuno-Horikawa, M. Nakano, M. Asahi, M. Takahashi, N. Uozumi, S. Ihara, S. H. Lee, Y. Ikeda, Y. Yamaguchi, Y. Aze, Y. Tomiyama, J. Fujii, K. Suzuki, A. Kondo,

- S. D. Shapiro, C. Lopez-Otin, T. Kuwaki, M. Okabe, K. Honke, and N. Taniguchi, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, 15791-6.
- [46] E. Al-Shareffi, J. L. Chaubard, C. Leonhard-Melief, S. K. Wang, C. H. Wong, and R. S. Haltiwanger, *Glycobiology* **2013**, 23, 188-98.
- [47] I. A. Wilson, J. J. Skehel, and D. C. Wiley, *Nature* **1981**, 289, 366-73.
- [48] D. C. Wiley, I. A. Wilson, and J. J. Skehel, *Nature* **1981**, 289, 373-8.
- [49] S. J. Watowich, J. J. Skehel, and D. C. Wiley, *Structure* **1994**, 2, 719-31.
- [50] P. M. Colman, J. N. Varghese, and W. G. Laver, *Nature* **1983**, 303, 41-4.
- [51] S. J. Crennell, E. F. Garman, W. G. Laver, E. R. Vimr, and G. L. Taylor, *Proc. Natl. Acad. Sci. U. S. A.* **1993**, 90, 9852-6.
- [52] W. Weis, J. H. Brown, S. Cusack, J. C. Paulson, J. J. Skehel, and D. C. Wiley, *Nature* **1988**, 333, 426-31.
- [53] D. C. Wiley and J. J. Skehel, *Ann. Rev. Biochem.* **1987**, 56, 365-94.
- [54] O. Blixt, S. Head, T. Mondala, C. Scanlan, M. E. Huflejt, R. Alvarez, M. C. Bryan, F. Fazio, D. Calarese, J. Stevens, N. Razi, D. J. Stevens, J. J. Skehel, I. van Die, D. R. Burton, I. A. Wilson, R. Cummings, N. Bovin, C. Wong, and J. C. Paulson, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101, 17033-8.
- [55] S. Hahn, *Nat. Struct. Mol. Biol.* **2004**, 11, 394-403.
- [56] E. W. Adams, D. M. Ratner, H. R. Bokesch, J. B. McMahon, B. R. O'Keefe, and P. H. Seeberger, *Chem. Biol.* **2004**, 11, 875-81.
- [57] Y. van Kooyk and T. B. H. Geijtenbeek, *Nat. Rev. Immunol.* **2003**, 3, 697-709.
- [58] R. T. Lee, T. L. Hsu, S. K. Huang, S. L. Hsieh, C. H. Wong, and Y. C. Lee, *Glycobiology* **2011**, 21, 512-20.

- [59] A. Holla and A. Skerra, *Protein Eng. Des. Sel.* **2011**, 24, 659-69.
- [60] L. de Witte, A. Nabatov, M. Pion, D. Fluitsma, M. A. W. P. de Jong, T. de Gruijl, V. Piguet, Y. van Kooyk, and T. B. H. Geijtenbeek, *Nat. Med.* **2007**, 13, 367-71.
- [61] Y. Guo, H. Feinberg, E. Conroy, D. A. Mitchell, R. Alvarez, O. Blixt, M. E. Taylor, W. I. Weis, and K. Drickamer, *Nat. Struct. Mol. Biol.* **2004**, 11, 591-8.
- [62] P-C. Pang, P. C. N. Chiu, C-L. Lee, L-Y. Chang, M. Panico, H. R. Morris, S. M. Haslam, K-H. Khoo, G. F. Clark, W. S. B. Yeung, and A. Dell, *Science* **2011**, 333, 1761-4.
- [63] Y. J. Kim and A. Varki, *Glycoconj. J.* **1997**, 14, 569-76.
- [64] T. Uchiyama, T. Woltering, W. Wong, C-C. Lin, T. Kajimoto, M. Takebayashi, G. Weitz-Schmidt, T. Asakura, M. Noda, and C-H. Wong, *Bioorg. Med. Chem.* **1996**, 4, 1149-65.
- [65] J. P. Abrahams, A. G. W. Leslie, R. Lutter, and John E. Walker, *Nature* **1994**, 370, 621-8.
- [66] P. Hensley, P. J. Mcdevitt, I. Brooks, J. J. Trill, J. A. Feild, D. E. McNulty, J. R. Connor, D. E. Griswold, N. V Kumar, K. D. Kopple, S. A. Carr, B. J. Dalton, and K. Johanson, *J. Biol. Chem.* **1994**, 269, 23949-58.
- [67] R. M. Cooke, R. S. Hale, S. G. Lister, G. Shah, and M. P. Weir, *Biochemistry* **1994**, 33, 10591-6.
- [68] G. S. Jacob, C. Kirmaier, S. Z. Abbas, S. C. Howard, C. N. Steininger, J. K. Welply, and P. Scudder, *Biochemistry* **1995**, 34, 1210-7.
- [69] M. P. Bevilacqua, S. Stengelin, M. A. Gimbrone, and B. Seed, *Science* **1989**, 243, 1160-5.

- [70] G. I. Johnston, R. G. Cook, and R. P. McEver, *Cell* **1989**, 56, 1033-44.
- [71] D. M. Lewinsohn, R. F. Bargatze, and E. C. Butcher, *J. Immunol.* **1987**, 138, 4313-21.
- [72] E. Larsen, A. Celi, G. E. Gilbert, B. C. Furie, J. K. Erban, R. Bonfanti, D. D. Wagner, and B. Furie, *Cell* **1989**, 59, 305-12.
- [73] W. A. Ducker, T. J. Senden, and R. M. Pashley, *Nature* **1990**, 343, 757-60.
- [74] M. P. Bevilacqua, J. S. Pober, D. L. Mendrick, R. S. Cotran, and M. A. Gimbrone, *Proc. Natl. Acad. Sci. U. S. A.* **1987**, 84, 9238-42.
- [75] E. L. Berg and M. K. Robinson, *J. Biol. Chem.* **1991**, 266, 14869-72.
- [76] L. M. Phillips, E. Nudelman, F. C. A. Gaeta, M. Perez, A. K. Singhai, S-I. Hakomori, and J. C. Paulson, *Science* **1990**, 250, 1130-2.
- [77] D. Tyrrell, P. James, N. Rao, C. Foxall, S. Abbas, F. Dasgupta, M. Nashed, A. Hasegawa, M. Kiso, and D. Asa, *Proc. Natl. Acad. Sci. U. S. A.* **1991**, 88, 10372-6.
- [78] G. Walz, A. Aruffo, W. Kolanus, M. Bevilacqua, and B. Seed, *Science* **1990**, 250, 1132-5.
- [79] C. T. Yuen, A. M. Lawson, W. Chai, M. Larkin, M. S. Stoll, A. C. Stuart, F. X. Sullivan, T. J. Ahern, and T. Feizi, *Biochemistry* **1992**, 31, 9126-31.
- [80] C-T. Yuen, K. Bezouska, J. O'Brien, M. Stoll, R. Lemoine, A. Lubineau, M. Kiso, A. Hasegawa, N. J. Bockovich, K. C. Nicolaou, and T. Feizi, *J. Biol. Chem.* **1994**, 269, 1595-8.
- [81] C. Foxall, S. R. Watson, D. Dowbenko, C. Fennie, L. A. Lasky, M. Kiso, A. Hasegawa, D. Asa, and B. K. Brandley, *J. Cell Biol.* **1992**, 117, 895-902.

- [82] E. L. Berg, J. Magnani, R. A. Warnock, M. K. Robinson, and E. C. Butcher, *Biochem. Biophys. Res. Commun.* **1992**, 184, 1048-55.
- [83] P. J. Green, T. Tamatani, T. Watanabe, M. Miyasaka, A. Hasegawa, M. Kiso, C-T. Yuen, M. S. Stoll, and T. Feizi, *Biochem. Biophys. Res. Commun.* **1992**, 188, 244-51.
- [84] M. J. Polley, M. L. Phillips, E. Wayner, E. Nudelmant, A. K. Singhalt, S. Hakomori, and J. C. P. Li, *Proc. Natl. Acad. Sci. U. S. A.* **1991**, 88, 6224-8.
- [85] E. Larsen, T. Palabrica, S. Sajer, G. E. Gilbert, D. D. Wagner, B. C. Furie, and B. Furie, *Cell* **1990**, 63, 467-74.
- [86] C. S. Shi, G. Y. Shi, S. M. Hsiao, Y. C. Kao, K. L. Kuo, C. Y. Ma, C. H. Kuo, B. I. Chang, C. F. Chang, C. H. Lin, C. H. Wong, and H. L. Wu, *Blood* **2008**, 112, 3661-70.
- [87] S. H. Rouhanifard, R. Xie, G. Zhang, X. Sun, X. Chen, and P. Wu, *Biomacromolecules* **2012**, 13, 3039-45.
- [88] R. Ignell et al., *Proc. Natl. Acad. Sci. U. S. A.* **2009**, 106, 18-23.
- [89] N. R. Sibson, A. M. Blamire, M. Bernades-Silva, S. Laurent, S. Boutry, R. N. Muller, P. Styles, and D. C. Anthony, *Magn. Reson. Med.* **2004**, 51, 248-52.
- [90] M. Bernardes-Silva, D. C. Anthony, A. C. Issekutz, and V. H. Perry, *J. Cereb. Blood Flow Metab.* **2001**, 21, 1115-24.
- [91] D. C. Anthony, S. J. Bolton, S. Fearn, and V. H. Perry, *Brain* **1997**, 120, 435-44.
- [92] A. Chrastina and J. E. Schnitzer, *Int. J. Nanomedicine* **2010**, 5, 653-9.

- [93] S. Y. Hong, G. Tobias, K. T. Al-Jamal, B. Ballesteros, H. Ali-Boucetta, S. Lozano-Perez, P. D. Nellist, R. B. Sim, C. Finucane, S. J. Mather, M. L. H. Green, K. Kostarelos, and B. G. Davis, *Nat. Mater.* **2010**, 9, 485-90.
- [94] Q. L. Lu, G. Bou-Gharios, and T. A. Partridge, *Gene Ther.* **2003**, 10, 131-42.
- [95] C. Borselli, H. Storrie, F. Benesch-Lee, D. Shvartsman, C. Cezar, J. W. Lichtman, H. H. Vandenburgh, and D. J. Mooney, *Proc. Natl. Acad. Sci.* **2010**, 107, 3287-92.
- [96] R. H. Weisbart, F. Yang, G. Chan, R. Wakelin, K. Ferreri, D. J. Zack, B. Harrison, L. A. Leinwand, and G. M. Cole, *Mol. Immunol.* **2003**, 39, 783-9.
- [97] Q. L. Lu, G. Bou-Gharios, and T. A. Partridge, *Gene Ther.* **2003**, 10, 131-42.
- [98] C. Borselli, H. Storrie, F. Benesch-Lee, D. Shvartsman, C. Cezar, J. W. Lichtman, H. H. Vandenburgh, and D. J. Mooney, *Proc. Natl. Acad. Sci.* **2010**, 107, 3287-92.
- [99] R. H. Weisbart, F. Yang, G. Chan, R. Wakelin, K. Ferreri, D. J. Zack, B. Harrison, L. A. Leinwand, and G. M. Cole, *Mol. Immunol.* **2003**, 39, 783-9.

6.7 Supporting information

General procedure for *in vivo* experiments

CD-1 male mice (20-25 g body weight) were injected *i.v.* with 0.15 mg of NPs (~4 mg nanoparticles/Kg mouse) in 100 μ L of saline solution (0.9 % NaCl in H₂O), under anaesthesia by isoflurane. The animals were left to recover. After 3 h mice were injected with 150-200 μ L of pentobarbital and perfused with saline solution until discolouration of the organs. Sections of liver (Li), lungs (Lu), heart (H), spleen (S), kidneys (K), gut (G), muscles (M1 from hind leg and M2 from front leg) and brain (B) were harvested for *ex-vivo* studies of the distribution of the particles.

All tests have been run in triplicate.

Two sections per organ were fixed in fixative solution (4% PFA) while other two sections were frozen in dry ice.

Fixative solution

PFA (40 g) was suspended in phosphate buffer (0.1 M, 1 L), and heated to 60-65 °C while stirring. To the suspension was added NaOH (1 M) dropwise, until the solution became clear. The solution was allowed to cool to r.t., adjusted to pH 7.2 with HCl 1 M and filtered.

Homogenised organs and counting

RIPA buffer

- 50 mM TRIS buffer, pH 8, with 150 mM NaCl
- 1% Triton X100
- 0.5% sodium deoxycholate
- 0.1% SDS

One frozen section for each organ was weighted, thawed to r.t., frozen again in liquid nitrogen and thawed once again in warm water. These sections were mechanically homogenised in RIPA buffer with 1 μ L of octanol. NPs were separated by magnet, washed (500 μ L) and recovered (30 μ L) with RIPA buffer. Prepared samples were counted on haemocytometer and the number of NPs per mg of tissue was calculated. Error bars refer to the mean SEM.

Dose studies

A set of mice were injected with GlcNAc-NPs (23% coverage, 4 mg/Kg); a second set was injected first with non-coated NPs and after 1 h with GlcNAc-NPs (23% coverage). Mice were culled after 3 h.

mouse	sample	sugar	coverage (%)
6-SDMB1	SDM625	GlcNAc	23
7-SDMB1	SDM625	GlcNAc	23
8-SDMB1	SDM625	GlcNAc	23
11-SDMB3	nanolink+SDM625	nanolink+GlcNAc	0+23
12-SDMB3	nanolink+SDM625	nanolink+GlcNAc	0+23
13-SDMB3	nanolink+SDM625	nanolink+GlcNAc	0+23

mass (mg)

liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
26.9	16.2	27.3	13.5	20.7	20.8	13.7	13.9	52.6
53.1	8.1	36.3	16.9	25.1	15.0	34.2	23.6	41.9
8.8	8.4	21.1	31.0	12.4	27.8	22.1	7.9	24.4
35.6	7.5	29.5	21.2	9.4	46.4	35.4	27.6	50.8
25.9	9.8	39.3	39.5	17	40.8	36.2	26.3	98.4
29.4	19	41.9	30.4	17	29.7	22.1	23.3	89.7

count (NPs per grid)

liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
2102	331	22	572	9	7	22	0	0
1710	103	12	282	3	15	17	1	0
2494	665	16	5130	15	12	23	1	1
2865	102	2	637	2	4	2	0	0
1581	103	3	473	1	3	2	0	1
2584	339	1	827	2	2	1	0	0

NPs/mg tissue

78	20	1	42	0	0	2	0	0
32	13	0	17	0	1	0	0	0
283	79	1	165	1	0	1	0	0
80	14	0	30	0	0	0	0	0
61	11	0	12	0	0	0	0	0
88	18	0	27	0	0	0	0	0

Distribution %

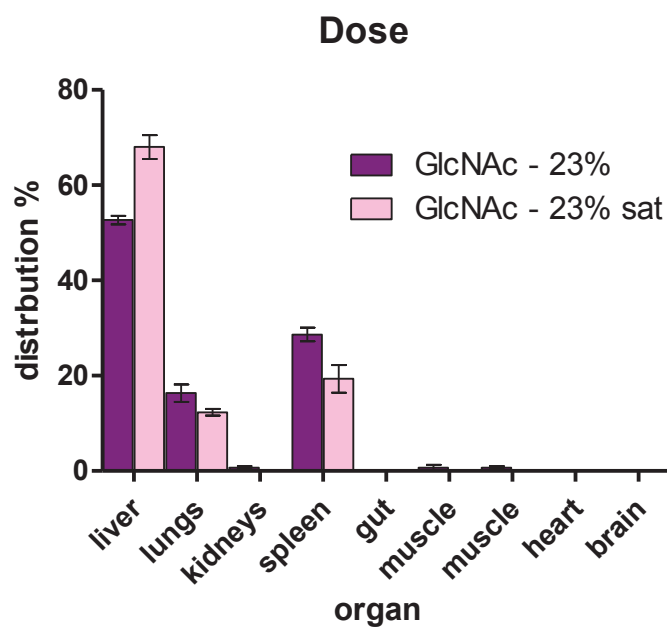
54	14	1	29	0	0	1	0	0
51	20	1	26	0	2	1	0	0
53	15	0	31	0	0	0	0	0
65	11	0	24	0	0	0	0	0
73	13	0	14	0	0	0	0	0
66	13	0	20	0	0	0	0	0

Distribution %

organ	GlcNAc - 23%			GlcNAc - 23% sat		
liver	54	51	53	65	73	66
lungs	14	20	15	11	13	13
kidneys	1	1	0	0	0	0
spleen	29	26	31	24	14	20
gut	0	0	0	0	0	0
muscle 1	0	2	0	0	0	0
muscle 2	1	1	0	0	0	0
heart	0	0	0	0	0	0
brain	0	0	0	0	0	0

AVG

organ	GlcNAc - 23%	GlcNAc - 23% sat
liver	53	68
lungs	16	12
kidneys	0	0
spleen	29	20
gut	0	0
muscle 1	1	0
muscle 2	1	0
heart	0	0
brain	0	0



(Fig. 6.2)

Coverage studies

Lac-NPs with 3 different coverage (14%, 35% and 68%) were injected in mice (4 mg/Kg) which were culled after 3 h.

mouse	sample	sugar	coverage (%)
2-SDMB1	SDM563	Lac	14
9-SDMB1	SDM563	Lac	14
10-SDMB1	SDM563	Lac	14
4-SDMB3	SDM670	Lac	35
5-SDMB3	SDM670	Lac	35
6-SDMB3	SDM670	Lac	35
1-SDMB3	SDM669	Lac	68
2-SDMB3	SDM669	Lac	68
3-SDMB3	SDM669	Lac	68

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
26.2	9.8	22.9	26.2	13.9	18.0	17.5	18.5	32.7
17.3	9.3	23.2	17.1	26.1	12.0	13.2	14.0	18.8
58.6	8.9	32.3	24.3	26.0	41.4	22.2	22.2	27.0
9.7	15.7	34.2	29.2	6.9	27.1	12.6	16.1	96.4
44.6	13.0	36.8	27.7	14.9	29.4	32	17.6	96.7
27.6	12.5	26.7	23.9	14.8	30.8	19.8	21.9	93.0
48.1	8.6	25.7	27.0	7.9	30.0	23.5	16.2	65.4
24.2	11.9	31.8	25.6	6.7	29.4	34.1	7.7	93.3
44.4	10.7	55.1	31.6	15.7	29.6	23.6	23.5	58.1

count (NPs per grid)								
425	4	1	17	3	65	44	1	0
475	10	8	19	7	90	92	11	0
333	5	5	17	0	46	37	2	1
80	9	20	56	3	48	33	5	3
142	8	3	14	1	52	48	3	7
210	7	13	16	1	57	63	0	2
401	9	15	32	8	108	112	8	5
317	3	4	22	5	128	131	2	0
343	4	1	11	4	104	48	0	0

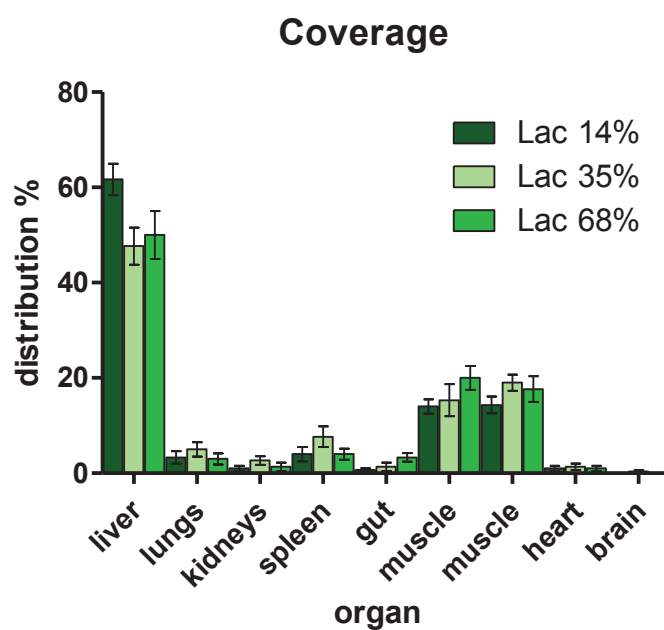
NPs/mg tissue								
16	0	0	1	0	4	3	0	0
27	1	0	1	0	8	7	1	0
6	1	0	1	0	1	2	0	0
8	1	1	2	0	2	3	0	0
3	1	0	1	0	2	2	0	0
8	1	0	1	0	2	3	0	0
8	1	1	1	1	4	5	0	0
13	0	0	1	1	4	4	0	0
8	0	0	0	0	4	2	0	0

Distribution %

68	2	0	3	1	15	11	0	0
60	2	1	2	1	16	15	2	0
57	6	2	7	0	11	17	1	0
50	3	4	12	3	11	16	2	0
40	8	1	6	1	22	19	2	1
53	4	3	5	0	13	22	0	0
40	5	3	6	5	17	23	2	0
56	1	1	4	3	18	16	1	0
54	3	0	2	2	25	14	0	0

organ	Lac - 14%			Lac - 35%			Lac - 68%		
liver	68	60	57	50	40	53	40	56	54
lungs	2	2	6	3	8	4	5	1	3
kidneys	0	1	2	4	1	3	3	1	0
spleen	3	2	7	12	6	5	6	4	2
gut	1	1	0	3	1	0	5	3	2
muscle 1	15	16	11	11	22	13	17	18	25
muscle 2	11	15	17	16	19	22	23	16	14
heart	0	2	1	2	2	0	2	1	0
brain	0	0	0	0	1	0	0	0	0

organ	AVG		
	Lac - 14%	Lac - 35%	Lac - 68%
liver	62	48	50
lungs	3	5	3
kidneys	1	3	1
spleen	4	8	4
gut	0	1	3
muscle 1	14	15	20
muscle 2	14	19	18
heart	1	1	1
brain	0	0	0



(Fig. 6.3)

Time studies

Lac-NPs (68% coverage, 4 mg/Kg) were tested, culling the animals after 1 h and 3 h from the injections.

mouse	sample	sugar	coverage (%)
5-SDMB2	SDM669	Lac	68
6-SDMB2	SDM669	Lac	68
7-SDMB2	SDM669	Lac	68
1-SDMB3	SDM669	Lac	68
2-SDMB3	SDM669	Lac	68
3-SDMB3	SDM669	Lac	68

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
39.0	11.5	31.0	17.4	10.1	24.8	29.5	27.8	30.3
23.9	10.4	23.6	24.0	12.6	15.6	28.6	18.1	46.3
22.5	8.7	40.1	22.3	17.8	18.3	31	19.2	31.6
48.1	8.6	25.7	27.0	7.9	30.0	23.5	16.2	65.4
24.2	11.9	31.8	25.6	6.7	29.4	34.1	7.7	93.3
44.4	10.7	55.1	31.6	15.7	29.6	23.6	23.5	58.1

count (NPs per grid)								
286	6	7	7	0	58	62	0	0
143	1	0	3	0	25	51	1	1
345	2	0	6	1	76	105	0	0
425	4	1	17	3	65	44	1	0
475	10	8	19	7	90	92	11	0
333	5	5	17	0	46	37	2	1

NPs/mg tissue								
7	1	0	0	0	2	2	0	0
6	0	0	0	0	2	2	0	0
15	0	0	0	0	4	3	0	0
9	0	0	1	0	2	2	0	0
20	1	0	1	1	3	3	1	0
8	0	0	1	0	2	2	0	0

Distribution %								
57	4	2	3	0	18	16	0	0
62	1	0	1	0	17	18	1	0
65	1	0	1	0	18	14	0	0
61	3	0	4	3	15	13	0	0
66	3	1	2	4	10	9	5	0
63	4	1	5	0	13	13	1	0

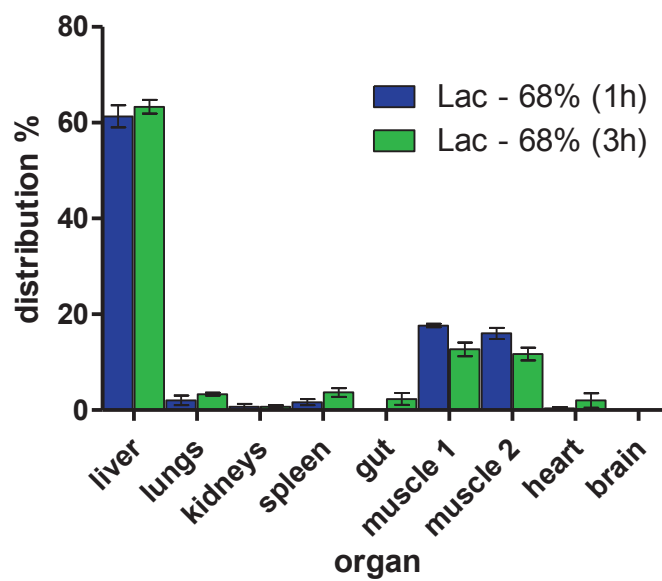
Distribution %

organ	Lac - 68% (1h)			Lac - 68% (3h)		
liver	57	62	65	61	66	63
lungs	4	1	1	3	3	4
kidneys	2	0	0	0	1	1
spleen	3	1	1	4	2	5
gut	0	0	0	3	4	0
muscle 1	18	17	18	15	10	13
muscle 2	16	18	14	13	9	13
heart	0	1	0	0	5	1
brain	0	0	0	0	0	0

AVG

organ	Lac - 68% (1h)	Lac - 68% (3h)
liver	61	64
lungs	2	3
kidneys	1	1
spleen	2	4
gut	0	2
muscle 1	17	13
muscle 2	16	12
heart	0	2
brain	0	0

Time



(Fig. 6.4)

Glycomap testing

NanoLink

mouse	sample	sugar	coverage (%)
3-SDMB1	NanoLink	-	0
4-SDMB1	NanoLink	-	0
5-SDMB1	NanoLink	-	0

mass (mg)

liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
16.2	8.0	16.2	19.7	42.9	23.8	22.7	18.0	57.7
17.6	15.6	21.7	24.6	22.7	19.5	20.1	17.6	24.4
22.4	17.8	38.2	18.4	18.8	9.5	20.2	25.2	74.1

count (NPs per grid)

462	0	3	332	3	9	0	0	0
257	6	0	235	1	3	2	3	1
470	2	8	330	4	4	0	7	0

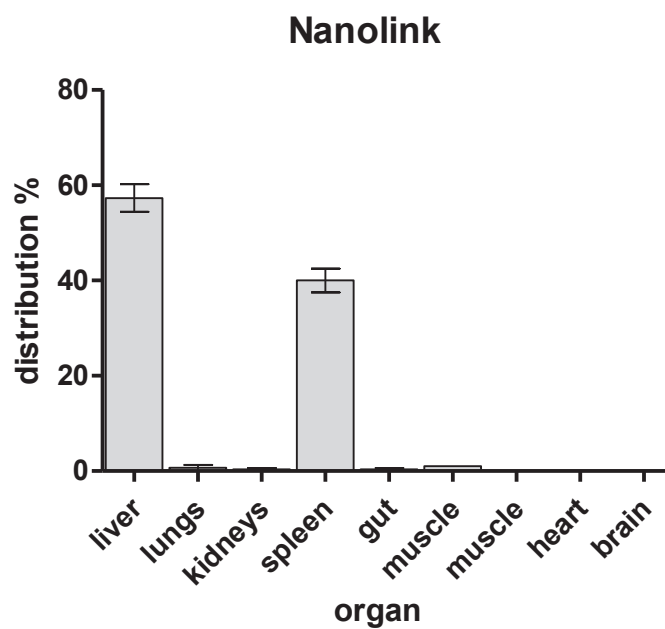
NPs/mg tissue

29	0	0	17	0	0	0	0	0
15	0	0	10	0	0	0	0	0
21	0	0	18	0	0	0	0	0

Distribution %

62	0	0	37	0	1	0	0	0
58	2	0	38	0	1	0	0	0
52	0	1	45	1	1	0	0	0

Distribution %				AVG	
organ	nanolink			organ	nanolink
liver	62	58	52	liver	58
lungs	0	2	0	lungs	1
kidneys	0	0	1	kidneys	0
spleen	37	38	45	spleen	40
gut	0	0	1	gut	0
muscle 1	1	1	1	muscle 1	1
muscle 2	0	0	0	muscle 2	0
heart	0	0	0	heart	0
brain	0	0	0	brain	0



(Fig. 6.1)

Glc

mouse	sample	sugar	coverage (%)
1-SDMB5	SDM630	Glc	70
2-SDMB5	SDM630	Glc	70
3-SDMB5	SDM630	Glc	70

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
39.0	13.4	39.7	50.7	19.6	37.7	28.7	30.4	42.6
104.4	24.5	66.9	19.0	15.0	66.7	30.4	23.6	109.3
21.0	10.0	56.1	20.0	21.6	20.3	32.5	25.6	69.2

count (NPs per grid)								
740	7	2	8	5	0	2	0	1
910	15	10	5	5	3	2	2	2
810	20	0	10	2	0	0	0	3

NPs/mg tissue								
19	1	0	0	0	0	0	0	0
9	1	0	0	0	0	0	0	0
39	2	0	1	0	0	0	0	0

Distribution %								
95	3	0	1	1	0	0	0	0
85	6	1	3	3	0	1	0	0
94	5	0	1	0	0	0	0	0

Distribution %					AVG	
organ	Glc			organ	Glc	
liver	95	85	94	liver	91	
lungs	3	6	5	lungs	4	
kidneys	0	1	0	kidneys	1	
spleen	1	3	1	spleen	2	
gut	1	3	0	gut	2	
muscle 1	0	0	0	muscle 1	0	
muscle 2	0	1	0	muscle 2	0	
heart	0	0	0	heart	0	
brain	0	0	0	brain	0	

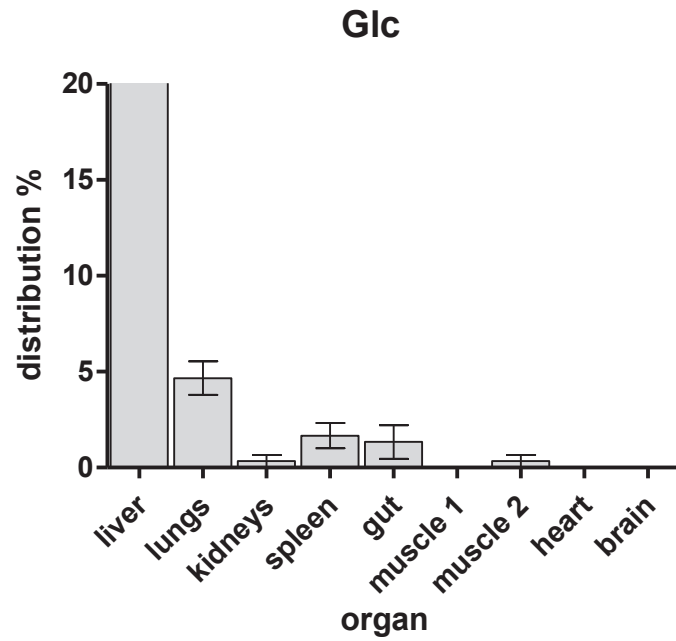


Figure S1 Glc-NPs distribution in CD-1 mice.

GlcNAc

mouse	sample	sugar	coverage (%)
4-SDMB5	SDM693	GlcNAc	68
5-SDMB5	SDM693	GlcNAc	68
6-SDMB5	SDM693	GlcNAc	68

mass (mg)

liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
15.9	24.8	56.8	29.8	15.8	27.7	31.0	26.8	56.6
21.6	23.0	62.6	18.5	13.8	12.8	13.5	45.8	68.1
22.8	10.9	21.3	26.0	13.6	29.7	22.8	27.7	83.2

count (NPs per grid)

308	87	12	44	0	8	4	0	36
300	77	53	38	0	4	6	4	0
318	43	10	19	1	4	0	0	9

NPs/mg tissue								
19	4	0	1	0	0	0	0	1
14	3	1	2	0	0	0	0	0
14	4	0	1	0	0	0	0	0

Distribution %								
76	14	1	6	0	1	1	0	2
67	16	4	10	0	2	1	0	0
72	20	2	4	0	1	0	0	1

Distribution %				AVG	
organ	GlcNAc			organ	GlcNAc
liver	76	67	72	liver	72
lungs	14	16	20	lungs	17
kidneys	1	4	2	kidneys	2
spleen	6	10	4	spleen	6
gut	0	0	0	gut	0
muscle 1	1	2	1	muscle 1	1
muscle 2	1	1	0	muscle 2	0
heart	0	0	0	heart	0
brain	2	0	1	brain	1

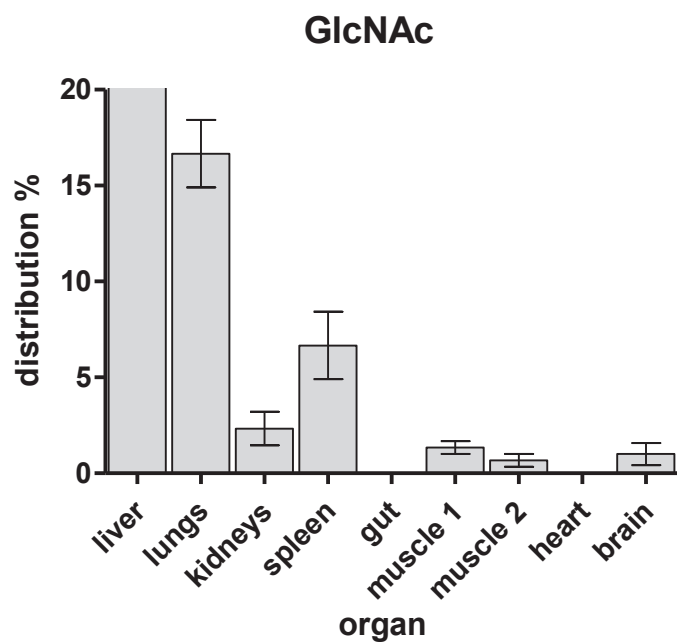


Figure S2 GlcNAc-NPs distribution in CD-1 mice.

Gal

mouse	sample	sugar	coverage (%)

1-SDMB6	SDM693	Gal	58
2-SDMB6	SDM693	Gal	58
3-SDMB6	SDM693	Gal	58

mass (mg)

liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
61.5	9.0	52.3	20.4	15.0	7.0	12.0	15.2	68.6
20.4	12.4	38.7	21.1	8.1	32.1	24.7	32.9	61.5
19.3	35.4	48.1	27.1	8.3	31.0	14.0	18.2	58.9

count (NPs per grid)

1200	0	0	22	1	3	2	2	0
947	2	31	59	1	9	0	0	2
861	1	2	156	1	0	0	0	7

NPs/mg tissue

20	0	0	1	0	0	0	0	0
46	0	1	3	0	0	0	0	0
45	0	0	6	0	0	0	0	0

Distribution %

92	0	0	5	0	2	1	0	0
92	0	2	6	0	1	0	0	0
88	0	0	11	0	0	0	0	0

Distribution %				AVG	
organ		Gal		organ	Gal
liver	92	92	88	liver	91
lungs	0	0	0	lungs	0
kidneys	0	2	0	kidneys	1
spleen	5	6	11	spleen	7
gut	0	0	0	gut	0
muscle 1	2	1	0	muscle 1	1
muscle 2	1	0	0	muscle 2	0
heart	0	0	0	heart	0
brain	0	0	0	brain	0

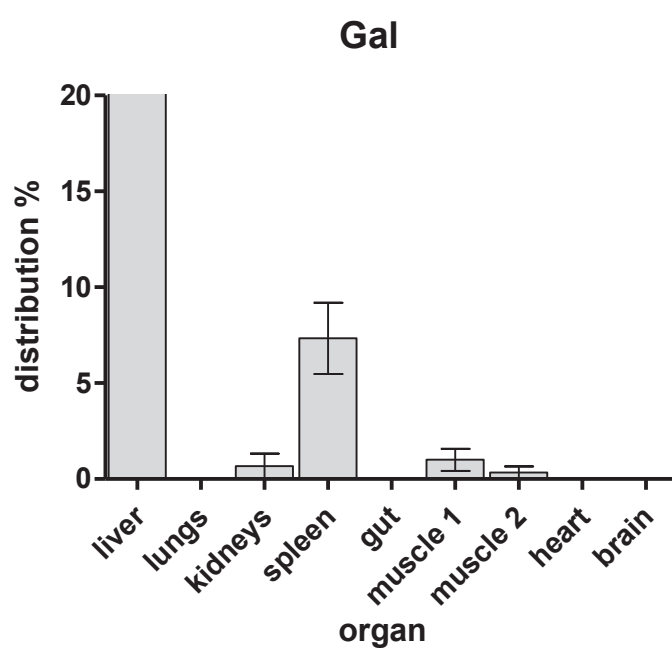


Figure S3 Gal-NPs distribution in CD-1 mice.

GalNAc

mouse	sample	sugar	coverage (%)
7-SDMB5	SDM649	GalNAc	67
8-SDMB5	SDM649	GalNAc	67
9-SDMB5	SDM649	GalNAc	67

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
15.5	8.6	27.2	15.3	8.4	35.9	33.6	20.6	67.6
28.3	18.1	38.5	14.3	16.5	18.2	18.1	28.2	72.0
40.3	21.3	22.1	22.8	9.3	30.7	33.4	33.4	83.8

count (NPs per grid)								
160	1	6	17	1	7	8	0	3
480	6	4	126	2	0	0	0	1
723	16	6	118	0	11	9	31	0

NPs/mg tissue								
10	0	0	1	0	0	0	0	0
17	0	0	9	0	0	0	0	0
18	1	0	5	0	0	0	0	0

Distribution %								
82	1	2	9	1	2	3	0	0
64	1	0	33	0	0	0	0	0
71	3	1	21	0	1	1	1	0

Distribution %				AVG	
organ		GalNAc		organ	GalNAc
liver	82	64	71	liver	73
lungs	1	1	3	lungs	2
kidneys	2	0	1	kidneys	1
spleen	9	33	21	spleen	21
gut	1	0	0	gut	0
muscle 1	2	0	1	muscle 1	1
muscle 2	3	0	1	muscle 2	1
heart	0	0	1	heart	0
brain	0	0	0	brain	0

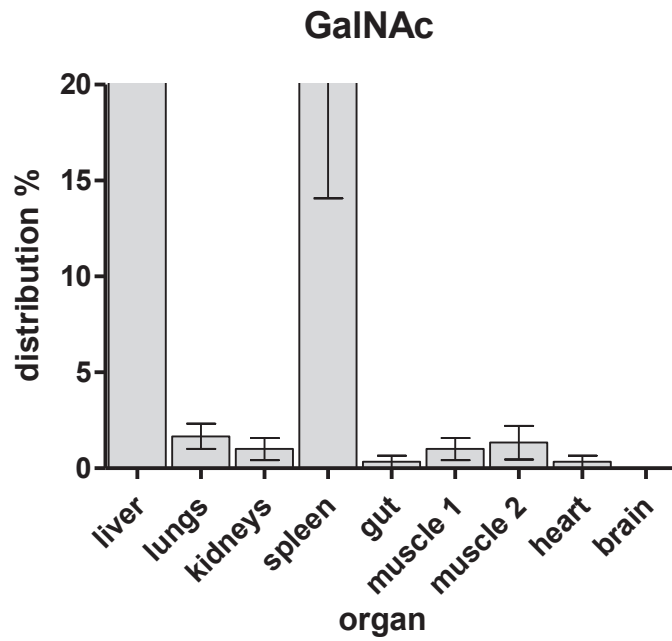


Figure S4 GalNAc-NPs distribution in CD-1 mice.

β Man

mouse	sample	sugar	coverage (%)
4-SDMB6	SDM693	bMan	62
5-SDMB6	SDM693	bMan	62
6-SDMB6	SDM693	bMan	62

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
33.7	21.1	68.7	29.1	11.1	27.2	28.8	25.3	48.5
39.1	10.3	58.5	28.5	9.7	19.6	18.6	38.4	65.5
13.6	19.9	73.8	26.8	11.7	30.4	19.5	41.1	75.7

count (NPs per grid)								
1124	15	0	165	0	0	0	4	5
979	16	0	29	1	1	0	2	2
215	25	5	19	5	0	0	2	1

NPs/mg tissue								
33	1	0	6	0	0	0	0	0
25	2	0	1	0	0	0	0	0
16	1	0	1	0	0	0	0	0

Distribution %								
84	2	0	14	0	0	0	0	0
90	6	0	4	0	0	0	0	0
86	7	0	4	2	0	0	0	0

Distribution %				AVG	
organ	bMan			organ	bMan
liver	84	90	86	liver	87
lungs	2	6	7	lungs	5
kidneys	0	0	0	kidneys	0
spleen	14	4	4	spleen	7
gut	0	0	2	gut	1
muscle 1	0	0	0	muscle 1	0
muscle 2	0	0	0	muscle 2	0
heart	0	0	0	heart	0
brain	0	0	0	brain	0

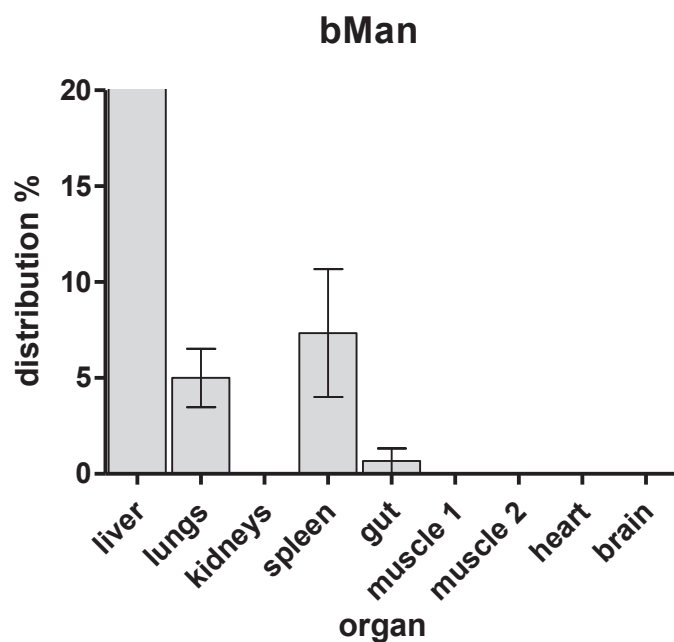


Figure S5 β Man-NPs distribution in CD-1 mice.

αMan

mouse	sample	sugar	coverage (%)
10-SDMB5	SDM631	aMan	59
11-SDMB5	SDM631	aMan	59
12-SDMB5	SDM631	aMan	59

mass (mg)

liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
19.3	13.1	48.4	29.1	16.4	21.6	23.5	20.0	100.8
25.9	24.0	48.7	32.5	21.3	16.2	9.8	29.1	87.0
14.6	11.7	35.4	20.0	7.7	32.4	28.7	39.0	62.0

count (NPs per grid)

565	33	2	17	1	1	0	1	4
9725	10	3	8	1	1	0	1	0
675	27	5	1	0	1	1	1	0

NPs/mg tissue

29	3	0	1	0	0	0	0	0
375	0	0	0	0	0	0	0	0
46	2	0	0	0	0	0	0	0

Distribution %

90	8	0	2	0	0	0	0	0
100	0	0	0	0	0	0	0	0
95	5	0	0	0	0	0	0	0

Distribution %				AVG	
organ		aMan		organ	aMan
liver	90	100	95	liver	95
lungs	8	0	5	lungs	4
kidneys	0	0	0	kidneys	0
spleen	2	0	0	spleen	1
gut	0	0	0	gut	0
muscle 1	0	0	0	muscle 1	0
muscle 2	0	0	0	muscle 2	0
heart	0	0	0	heart	0
brain	0	0	0	brain	0

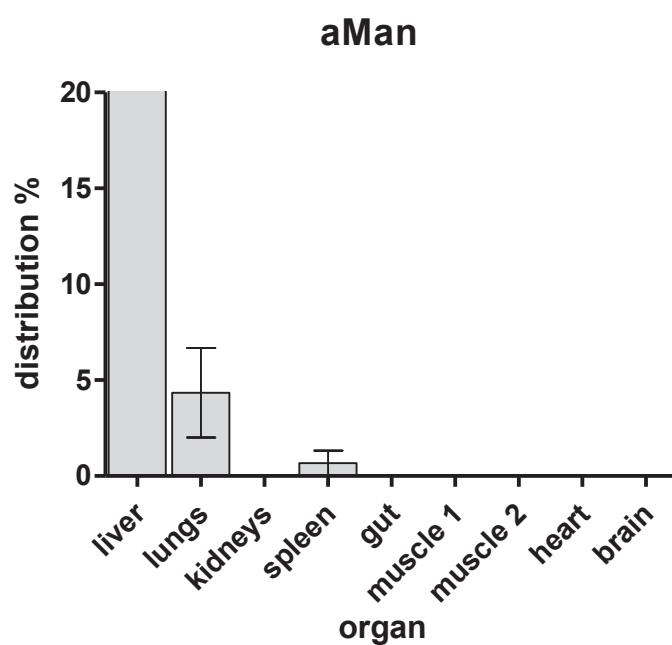


Figure S6 αMan-NPs distribution in CD-1 mice.

ManNAc

mouse	sample	sugar	coverage (%)
7-SDMB6	SDM693	ManNAc	60
8-SDMB6	SDM693	ManNAc	60
9-SDMB6	SDM693	ManNAc	60

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
54.0	6.3	49.7	21.2	9.5	7.9	12.9	37.1	109.0
30.0	14.8	34.1	23.9	9.7	22.7	22.5	12.7	66.9
30.0	46.6	26.7	31.5	7.9	30.8	14.5	28.7	99.4

count (NPs per grid)								
596	5	47	23	6	0	1	4	2
1049	9	0	30	2	28	0	1	0
1256	4	4	48	1	24	0	8	2

NPs/mg tissue								
11	1	1	1	1	0	0	0	0
35	1	0	1	0	1	0	0	0
42	0	0	2	0	1	0	0	0

Distribution %								
76	5	6	7	4	0	0	0	0
91	2	0	3	1	3	0	0	0
94	0	0	3	0	2	0	0	0

Distribution %				AVG	
organ		ManNAc		organ	ManNAc
liver	76	91	94	liver	87
lungs	5	2	0	lungs	2
kidneys	6	0	0	kidneys	2
spleen	7	3	3	spleen	5
gut	4	1	0	gut	2
muscle 1	0	3	2	muscle 1	2
muscle 2	0	0	0	muscle 2	0
heart	0	0	0	heart	0
brain	0	0	0	brain	0

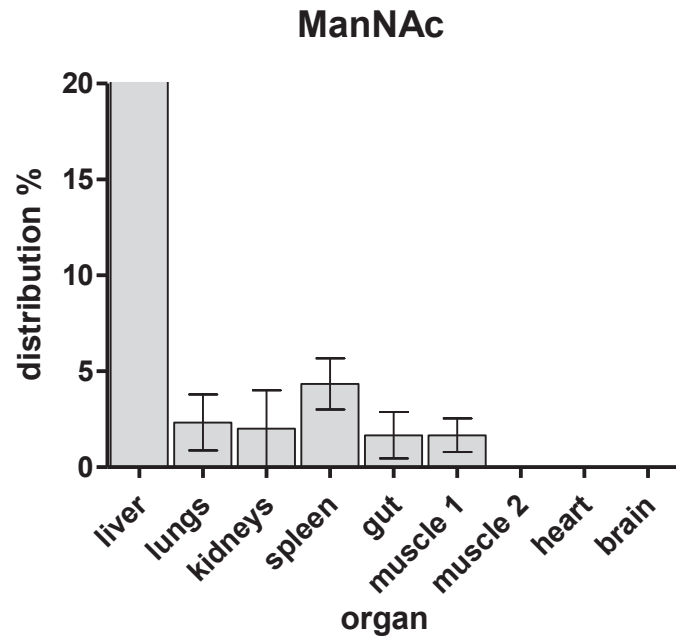


Figure S7 ManNAc-NPs distribution in CD-1 mice.

Xyl

mouse	sample	sugar	coverage (%)
10-SDMB6	SDM693	Xyl	57
11-SDMB6	SDM693	Xyl	57
12-SDMB6	SDM693	Xyl	57

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
55.8	12.6	42.4	39.5	11.3	13.3	22.1	8.7	30.2
31.9	13.3	44.6	28.0	12.8	20.1	23.1	18.4	73.7
20.9	16.1	51.8	22.6	5.9	24.8	8.7	21.8	85.4

count (NPs per grid)								
586	9	18	210	7	2	0	2	0
267	3	12	189	4	1	0	3	7
77	17	15	77	0	1	5	1	0

NPs/mg tissue								
11	1	0	5	1	0	0	0	0
8	0	0	7	0	0	0	0	0
4	1	0	3	0	0	0	0	0

Distribution %								
59	4	2	30	3	1	0	0	0
52	1	2	42	2	0	0	0	1
42	12	3	39	0	0	3	0	0

Distribution %				AVG	
organ		Xyl		organ	Xyl
liver	59	52	42	liver	51
lungs	4	1	12	lungs	6
kidneys	2	2	3	kidneys	2
spleen	30	42	39	spleen	37
gut	3	2	0	gut	2
muscle 1	1	0	0	muscle 1	1
muscle 2	0	0	3	muscle 2	1
heart	0	0	0	heart	0
brain	0	1	0	brain	0

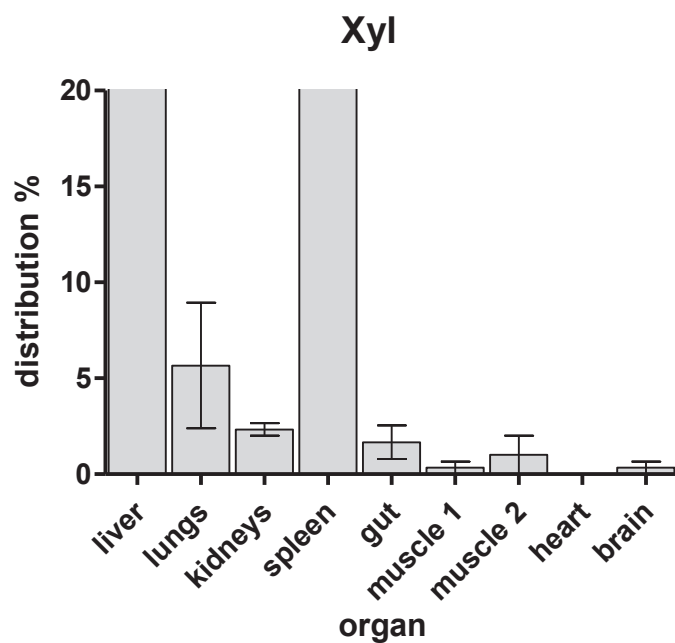


Figure S8 Xyl-NPs distribution in CD-1 mice.

Fuc

		mouse	sample	sugar	coverage (%)			
		13-SDMB6	SDM693	Fuc	65			
		14-SDMB6	SDM693	Fuc	65			
		15-SDMB6	SDM693	Fuc	65			
mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
26.0	7.0	73.3	13.6	14.3	16.2	18.5	28.6	54.9
65.8	14.0	41.6	19.5	6.3	21.6	11.3	19.4	57.0
29.2	21.7	42.0	29.6	6.1	22.4	19.9	17.3	35.4

count (NPs per grid)								
460	12	1	507	1	0	21	1	10
290	2	0	127	0	3	0	0	1
360	36	7	398	1	1	0	0	3

NPs/mg tissue								
18	2	0	37	0	0	1	0	0
4	0	0	7	0	0	0	0	0
12	2	0	13	0	0	0	0	0

Distribution %								
31	3	0	65	0	0	1	0	0
39	1	0	58	0	1	0	0	0
44	6	1	48	1	0	0	0	0

Distribution %				AVG	
organ		Fuc		organ	Fuc
liver	31	39	44	liver	38
lungs	3	1	6	lungs	3
kidneys	0	0	1	kidneys	0
spleen	65	58	48	spleen	57
gut	0	0	1	gut	0
muscle 1	0	1	0	muscle 1	0
muscle 2	1	0	0	muscle 2	0
heart	0	0	0	heart	0
brain	0	0	0	brain	0

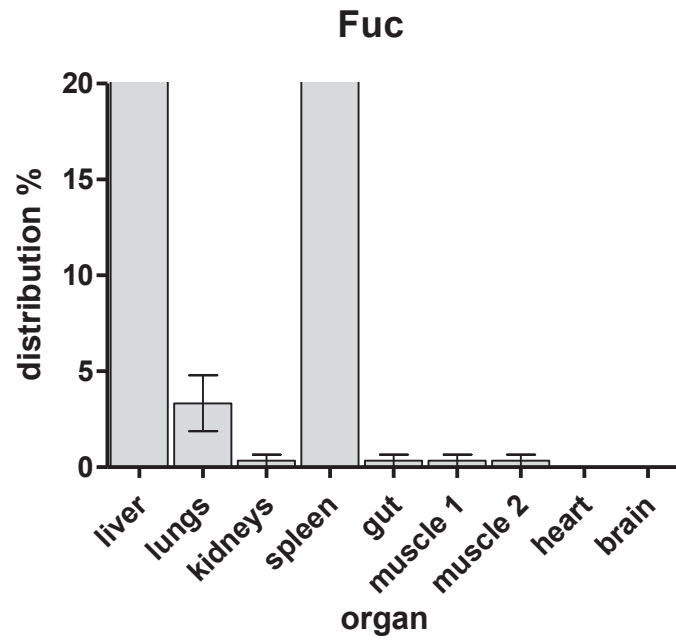


Figure S9 Fuc-NPs distribution in CD-1 mice.

GlcA

mouse	sample	sugar	coverage (%)
1-SDMB7	SDM695	GlcA	70
2-SDMB7	SDM695	GlcA	70
3-SDMB7	SDM695	GlcA	70

mass (mg)

liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
27.1	13.2	28.4	17.3	5.6	11.9	5.6	23.5	75.2
44.3	15.4	32.1	33.6	11.6	10.2	14.2	31.2	102.2
15.3	17.1	44.2	48.4	16.0	24.0	18.3	27.0	72.0

count (NPs per grid)

1277	0	2	34	0	3	0	3	0
922	1	0	56	0	0	0	2	2
1960	4	0	102	2	3	0	1	0

NPs/mg tissue								
47	0	0	2	0	0	0	0	0
21	0	0	2	0	0	0	0	0
128	0	0	2	0	0	0	0	0

Distribution %								
95	0	0	4	0	1	0	0	0
92	0	0	7	0	0	0	0	0
98	0	0	2	0	0	0	0	0

Distribution %				AVG	
organ	GlcA			organ	GlcA
liver	95	92	98	liver	95
lungs	0	0	0	lungs	0
kidneys	0	0	0	kidneys	0
spleen	4	7	2	spleen	4
gut	0	0	0	gut	0
muscle 1	1	0	0	muscle 1	0
muscle 2	0	0	0	muscle 2	0
heart	0	0	0	heart	0
brain	0	0	0	brain	0

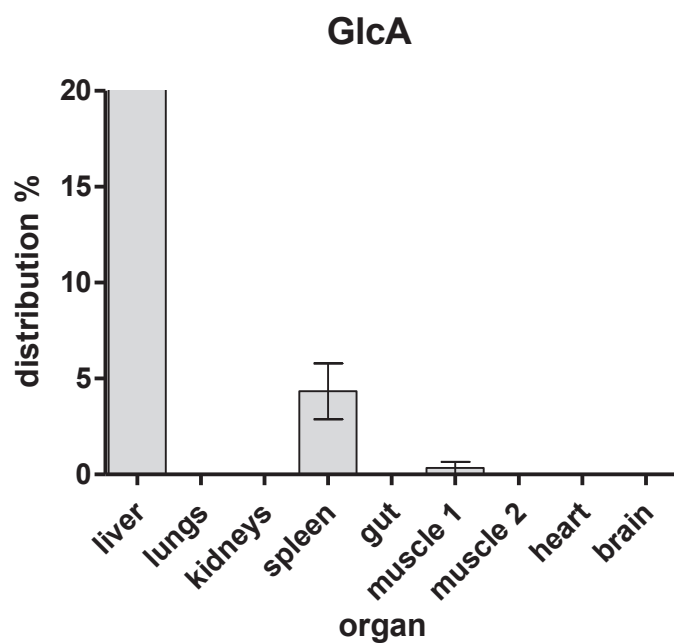


Figure S10 GlcA-NPs distribution in CD-1 mice.
402

Neu

mouse	sample	sugar	coverage (%)

4-SDMB7	SDM695	Neu	66
5-SDMB7	SDM695	Neu	66
6-SDMB7	SDM695	Neu	66

mass (mg)

liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
37.2	19.4	35.3	29.7	20.2	22.0	20.2	14.9	53.2
10.1	12.1	19.8	15.7	4.8	17.3	13.7	16.5	58.5
24.9	16.8	13.6	11.5	5.5	6.5	10.4	19.5	79.2

count (NPs per grid)

712	9	1	2	1	2	6	0	0
632	3	0	77	0	5	0	1	0
500	12	15	0	0	1	0	1	0

NPs/mg tissue

19	0	0	0	0	0	0	0	0
63	0	0	5	0	0	0	0	0
20	1	1	0	0	0	0	0	0

Distribution %

95	2	0	0	0	0	2	0	0
92	0	0	7	0	0	0	0	0
91	3	5	0	0	1	0	0	0

Distribution %				AVG	
organ	Neu			organ	Neu
liver	95	92	91	liver	93
lungs	2	0	3	lungs	2
kidneys	0	0	5	kidneys	2
spleen	0	7	0	spleen	3
gut	0	0	0	gut	0
muscle 1	0	0	1	muscle 1	1
muscle 2	2	0	0	muscle 2	1
heart	0	0	0	heart	0
brain	0	0	0	brain	0

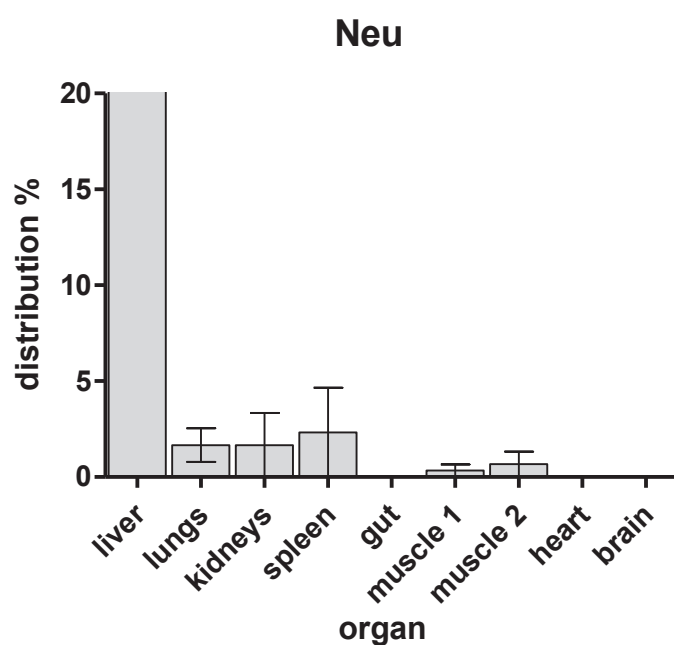


Figure S11 Neu-NPs distribution in CD-1 mice.

Lac

mouse	sample	sugar	coverage (%)
1-SDMB3	SDM669	Lac	68
2-SDMB3	SDM669	Lac	68
3-SDMB3	SDM669	Lac	68

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
48.1	8.6	25.7	27.0	7.9	30.0	23.5	16.2	65.4
24.2	11.9	31.8	25.6	6.7	29.4	34.1	7.7	93.3
44.4	10.7	55.1	31.6	15.7	29.6	23.6	23.5	58.1

count (NPs per grid)								
425	4	1	17	3	65	44	1	0
475	10	8	19	7	90	92	11	0
333	5	5	17	0	46	37	2	1

NPs/mg tissue								
9	0	0	1	0	2	3	0	0
20	1	0	1	1	3	12	0	0
8	0	0	1	0	2	2	0	0

Distribution %								
58	3	0	4	2	14	18	0	0
52	2	1	2	3	8	32	0	0
64	4	1	5	0	13	13	0	0

Distribution %				AVG	
organ		Lac		organ	Lac
liver	58	52	64	liver	58
lungs	3	2	4	lungs	3
kidneys	0	1	1	kidneys	1
spleen	4	2	5	spleen	4
gut	2	3	0	gut	2
muscle 1	14	8	13	muscle 1	12
muscle 2	18	32	13	muscle 2	21
heart	0	0	0	heart	0
brain	0	0	0	brain	0

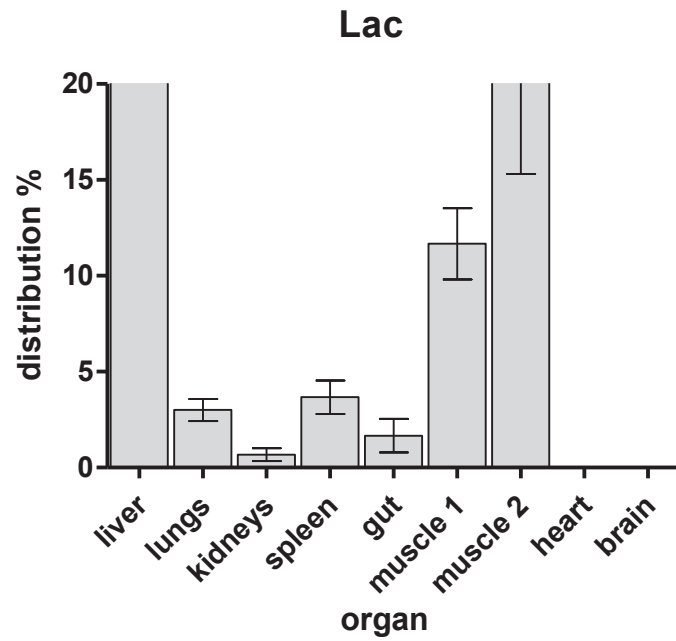


Figure S12 Lac-NPs distribution in CD-1 mice.

(1,4)-LacNAc

mouse	sample	sugar	coverage (%)
7-SDMB7	SDM695	(1,4)-LacNAc	62
8-SDMB7	SDM695	(1,4)-LacNAc	62
9-SDMB7	SDM695	(1,4)-LacNAc	62

mass (mg)

liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
42.4	9.6	37.4	23.0	6.4	15.3	12.8	7.1	64.6
17.6	5.1	24.3	17.0	7.1	8.0	15.2	12.6	54.0
13.9	12.6	30.3	17.2	5.8	20.5	19.0	24.3	56.3

count (NPs per grid)

504	1	1	17	0	3	0	0	1
620	1	6	40	0	6	1	2	0
107	2	1	3	1	0	1	1	0

NPs/mg tissue								
12	0	0	1	0	0	0	0	0
35	0	0	2	0	1	0	0	0
8	0	0	0	0	0	0	0	0

Distribution %								
92	1	0	6	0	2	0	0	0
91	1	1	6	0	2	0	0	0
93	2	0	2	2	0	0	0	0

Distribution %			AVG	
organ	(1,4)-LacNac			(1,4)-LacNac
liver	92	91	93	92
lungs	1	1	2	1
kidneys	0	1	0	0
spleen	6	6	2	5
gut	0	0	2	1
muscle 1	2	2	0	1
muscle 2	0	0	0	0
heart	0	0	0	0
brain	0	0	0	0

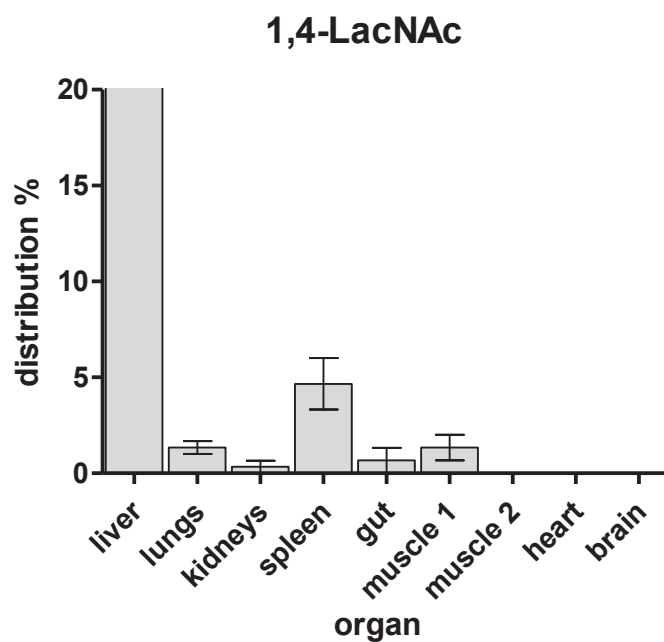


Figure S13 1-4,LacNac-NPs distribution in CD-1 mice.

(1,3)-LacNAc

		mouse	sample	sugar		coverage (%)		
		10-SDMB7	SDM695	(1,3)-LacNAc		55		
		11-SDMB7	SDM695	(1,3)-LacNAc		55		
		12-SDMB7	SDM695	(1,3)-LacNAc		55		
mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
43.8	6.7	34.7	18.7	7.4	17.1	16.8	16.5	59.7
23.4	17.2	24.6	18.3	5.9	16.8	11.6	19.9	79.0
9.9	9.9	14.2	24.5	4.2	11.2	7.9	17.2	90.2
count (NPs per grid)								
181	1	0	46	2	11	3	6	0
100	2	0	41	0	13	0	0	0
60	10	0	45	0	14	1	1	0
NPs/mg tissue								
4	0	0	2	0	1	0	0	0
4	0	0	2	0	1	0	0	0
6	1	0	2	0	1	0	0	0
Distribution %								
52	2	0	31	3	8	2	1	0
58	2	0	30	0	10	0	0	0
59	10	0	18	0	12	1	0	0

Distribution %				AVG	
organ	(1,3)-LacNAc			organ	(1,3)-LacNAc
liver	52	58	59	liver	56
lungs	2	2	10	lungs	4
kidneys	0	0	0	kidneys	0
spleen	31	30	18	spleen	26
gut	3	0	0	gut	1
muscle 1	8	10	12	muscle 1	10
muscle 2	2	0	1	muscle 2	1
heart	1	0	0	heart	0
brain	0	0	0	brain	0

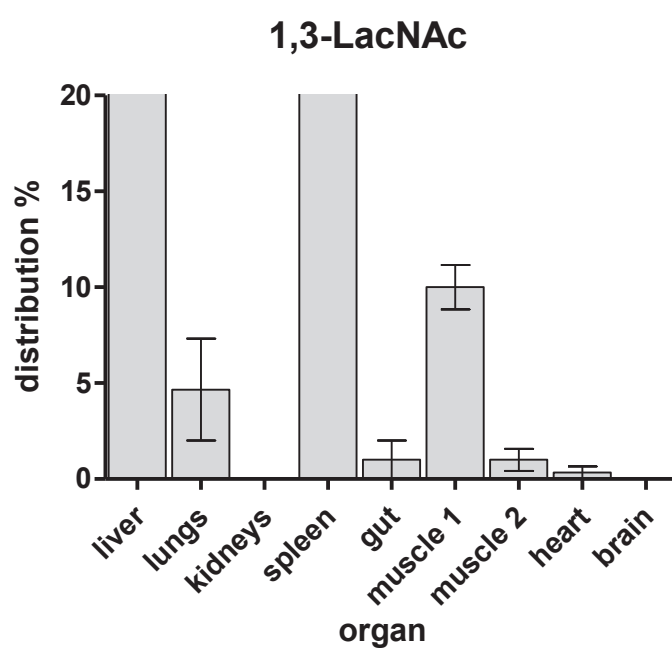


Figure S14 1-3,LacNAc-NPs distribution in CD-1 mice.

sLacNAc

mouse	sample	sugar	coverage (%)
13-SDMB7	SDM695	sLacNAc	67
14-SDMB7	SDM695	sLacNAc	67
15-SDMB7	SDM695	sLacNAc	67

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
19.8	20.9	21.3	15.0	8.1	9.3	9.4	17.9	55.5
7.9	13.1	21.5	12.6	5.4	7.7	9.8	35.5	68.4
12.1	8.0	23.7	15.1	4.4	16.9	9.6	16.6	59.6

count (NPs per grid)								
217	2	1	0	0	7	0	8	0
153	0	7	4	0	0	2	0	0
170	4	2	1	4	1	0	0	0

NPs/mg tissue								
11	0	0	0	0	1	0	0	0
19	0	0	0	0	0	0	0	0
14	1	0	0	1	0	0	0	0

Distribution %								
91	1	0	0	0	6	0	1	0
97	0	2	2	0	0	0	0	0
90	3	1	0	6	0	0	0	0

Distribution %				AVG	
organ	sLacNAc			organ	sLacNAc
liver	91	97	90	liver	93
lungs	1	0	3	lungs	1
kidneys	0	2	1	kidneys	1
spleen	0	2	0	spleen	1
gut	0	0	6	gut	2
muscle 1	6	0	0	muscle 1	2
muscle 2	0	0	0	muscle 2	0
heart	1	0	0	heart	0
brain	0	0	0	brain	0

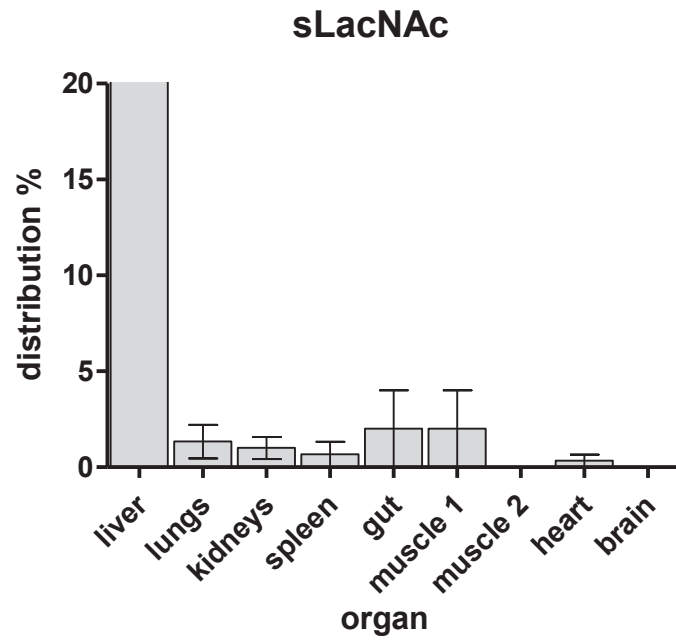


Figure S15 sLacNAc-NPs distribution in CD-1 mice.

sLe^x

mouse	sample	sugar	coverage (%)
1-SDMB8	SDM695	sLe ^x	71
2-SDMB8	SDM695	sLe ^x	71
3-SDMB8	SDM695	sLe ^x	71

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
26.9	16.4	35.1	11.5	3.2	9.4	12.6	10.9	97.0
26.9	20.9	45.5	12.5	8.0	18.6	9.9	28.2	21.9
24.3	23.4	57.4	20.4	12.2	21.0	8.4	21.5	88.3

count (NPs per grid)								
376	6	1	5	1	0	0	3	1
836	6	7	4	0	0	3	0	0

456	1	1	6	9	4	0	0	0

NPs/mg tissue

14	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0
19	0	0	0	1	0	0	0	0

Distribution %

92	2	0	3	2	0	0	0	0
97	1	0	1	0	0	0	0	0
94	0	0	1	4	1	0	0	0

Distribution %				AVG	
organ	sLex			organ	sLex
liver	92	97	94	liver	94
lungs	2	1	0	lungs	1
kidneys	0	0	0	kidneys	0
spleen	3	1	1	spleen	2
gut	2	0	4	gut	2
muscle 1	0	0	1	muscle 1	0
muscle 2	0	0	0	muscle 2	0
heart	0	0	0	heart	0
brain	0	0	0	brain	0

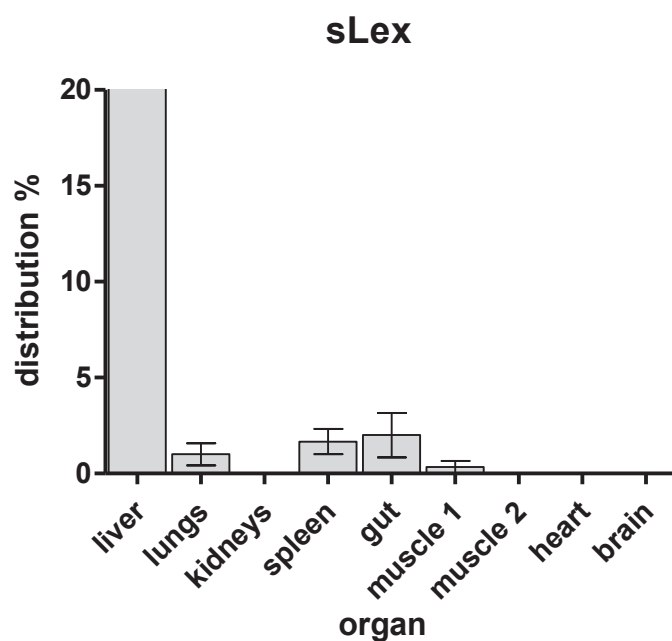


Figure S16 sLe^x-NPs distribution in CD-1 mice.
412

Le^x

mouse sample sugar coverage (%)

4-SDMB8	SDM695	Le ^x	66
5-SDMB8	SDM695	Le ^x	66
6-SDMB8	SDM695	Le ^x	66

mass (mg)

liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
12.6	8.7	31.1	11.6	9.1	18.4	14.5	30.3	70.4
9.6	19.5	40.6	23.8	8.4	4.8	28.9	24.9	55.7
41.1	40.6	38.5	18.6	15.0	18.2	15.4	44.3	72.7

count (NPs per grid)

157	1	0	2	0	2	0	0	0
560	1	14	45	2	0	0	0	0
875	12	0	4	1	0	0	1	0

NPs/mg tissue

12	0	0	0	0	0	0	0	0
58	0	0	2	0	0	0	0	0
21	0	0	0	0	0	0	0	0

Distribution %

97	1	0	1	0	1	0	0	0
96	0	1	3	0	0	0	0	0
97	1	0	1	0	0	0	0	0

Distribution %				AVG	
organ	Lex			organ	Lex
liver	97	96	97	liver	97
lungs	1	0	1	lungs	1
kidneys	0	1	0	kidneys	0
spleen	1	3	1	spleen	2
gut	0	0	0	gut	0
muscle 1	1	0	0	muscle 1	0
muscle 2	0	0	0	muscle 2	0
heart	0	0	0	heart	0
brain	0	0	0	brain	0

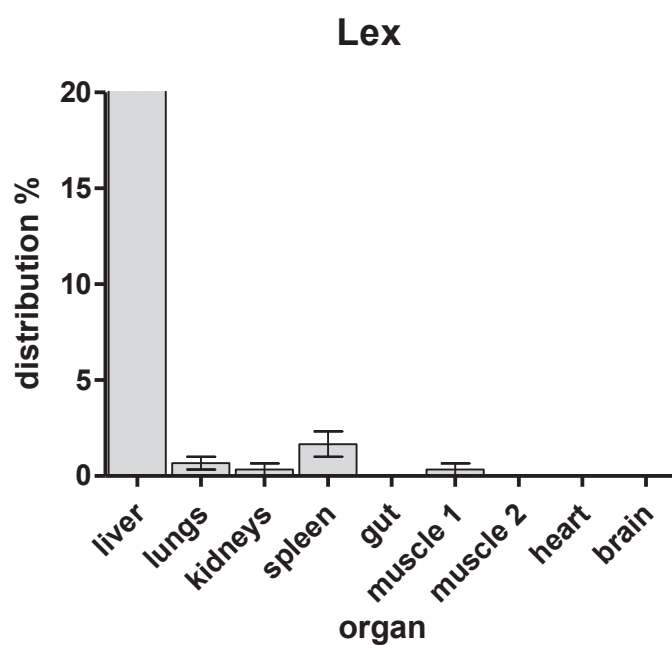


Figure S17 Le^x-NPs distribution in CD-1 mice.

Le^y

mouse	sample	sugar	coverage (%)
7-SDMB8	SDM695	Le ^y	60
8-SDMB8	SDM695	Le ^y	60
9-SDMB8	SDM695	Le ^y	60

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
47.4	9.2	38.2	19.7	16.1	19.1	6.2	31.1	112.4
19.2	35.4	43.9	16.3	12.8	29.4	10.1	25.3	61.8
18.2	27.0	37.9	15.8	12.9	9.6	12.7	20.6	95.5

count (NPs per grid)								
1225	1	0	14	0	0	1	6	0
455	1	1	33	1	2	2	1	0
303	4	4	62	0	3	0	0	1

NPs/mg tissue								
26	0	0	1	0	0	0	0	0
24	0	0	2	0	0	0	0	0
17	0	0	4	0	0	0	0	0

Distribution %								
97	0	0	3	0	0	0	0	0
91	0	0	8	0	0	0	0	0
79	1	0	19	0	1	0	0	0

Distribution %				AVG	
organ	Ley			organ	Ley
liver	97	91	79	liver	89
lungs	0	0	1	lungs	0
kidneys	0	0	0	kidneys	0
spleen	3	8	19	spleen	10
gut	0	0	0	gut	0
muscle 1	0	0	1	muscle 1	1
muscle 2	0	0	0	muscle 2	0
heart	0	0	0	heart	0
brain	0	0	0	brain	0

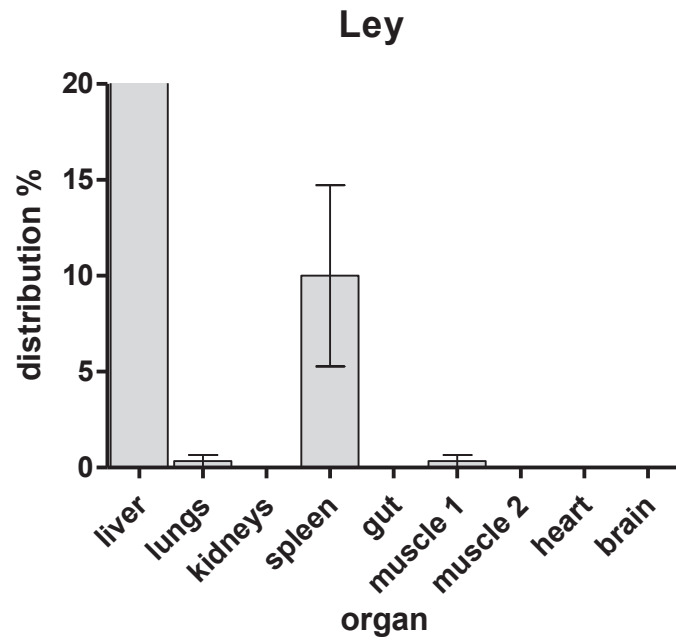


Figure S18 Le^Y-NPs distribution in CD-1 mice.

Le^a

mouse	sample	sugar	coverage (%)
10-SDMB8	SDM695	Le ^a	55
11-SDMB8	SDM695	Le ^a	55
12-SDMB8	SDM695	Le ^a	55

mass (mg)

liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
30.8	26.2	33.3	14.0	7.4	11.9	12.2	28.9	75.4
23.6	23.8	24.4	6.8	5.8	12.2	4.6	10.0	70.5
22.7	14.9	25.5	12.8	2.6	17.0	6.0	14.8	45.2

count (NPs per grid)

359	11	5	29	0	2	6	0	0
655	2	20	51	6	0	1	4	1
1162	3	44	105	3	39	0	0	0

NPs/mg tissue								
12	0	0	2	0	0	0	0	0
28	0	1	8	1	0	0	0	0
51	0	2	8	1	2	0	0	0

Distribution %								
79	3	1	14	0	1	1	0	0
74	0	2	20	3	0	0	0	0
79	0	3	13	2	4	0	0	0

Distribution %						AVG
organ	Lea			organ	Lea	
liver	79	74	79	liver	78	
lungs	3	0	0	lungs	1	
kidneys	1	2	3	kidneys	2	
spleen	14	20	13	spleen	16	
gut	0	3	2	gut	2	
muscle 1	1	0	4	muscle 1	2	
muscle 2	1	0	0	muscle 2	1	
heart	0	0	0	heart	0	
brain	0	0	0	brain	0	

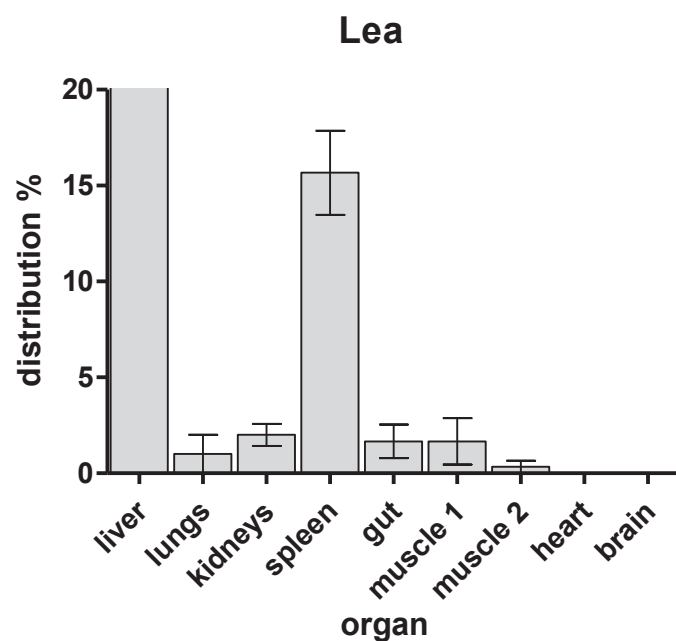


Figure S19 Le^a-NPs distribution in CD-1 mice.

Le^b

mouse	sample	sugar	coverage (%)
13-SDMB8	SDM695	Le ^b	56
14-SDMB8	SDM695	Le ^b	56
15-SDMB8	SDM695	Le ^b	56

mass (mg)								
liver	lungs	kidneys	spleen	gut	muscle1	muscle2	heart	brain
14.3	12.9	51.5	7.0	6.7	27.0	13.9	31.7	65.9
59.0	15.9	45.3	14.6	6.4	29.3	15.4	25.7	80.3
25.6	13.0	46.8	9.9	7.1	18.3	14.8	19.1	75.3

count (NPs per grid)								
845	2	10	209	7	0	0	10	0
410	0	2	37	3	1	2	3	0
280	9	1	35	1	0	0	2	0

NPs/mg tissue								
59	0	0	30	1	0	0	0	0
7	0	0	3	0	0	0	0	0
11	1	0	4	0	0	0	0	0

Distribution %								
65	0	0	33	1	0	0	0	0
68	0	0	25	5	0	1	0	0
71	5	0	23	1	0	0	0	0

Distribution %				AVG	
organ		Leb		organ	Leb
liver	65	68	71	liver	68
lungs	0	0	5	lungs	2
kidneys	0	0	0	kidneys	0
spleen	33	25	23	spleen	27
gut	1	5	1	gut	2
muscle 1	0	0	0	muscle 1	0
muscle 2	0	1	0	muscle 2	0
heart	0	0	0	heart	0
brain	0	0	0	brain	0

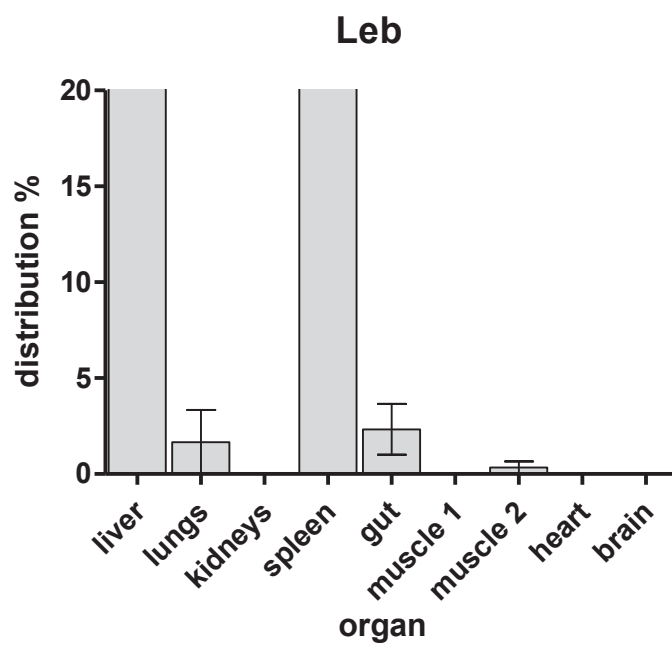
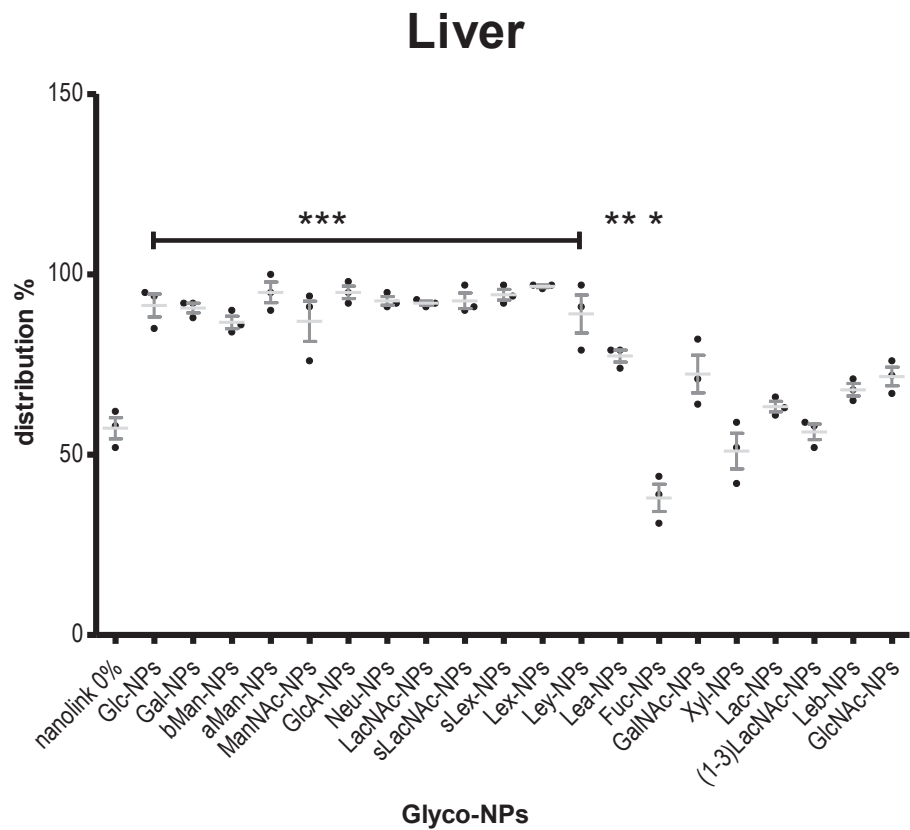
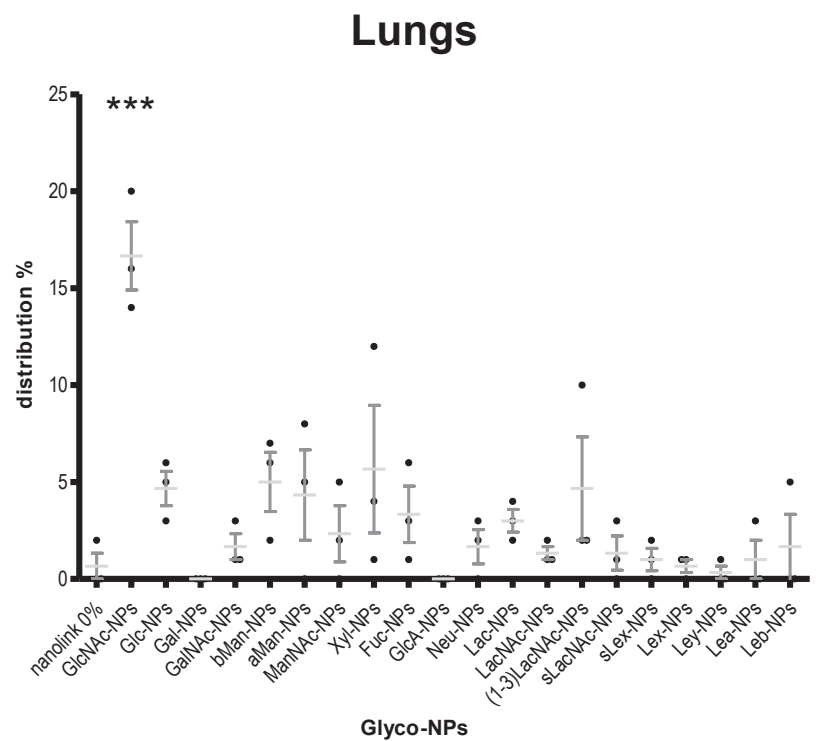


Figure S20 Leb^b-NPs distribution in CD-1 mice.

Statistical studies



(Figure 6.5)



(Figure 6.7)

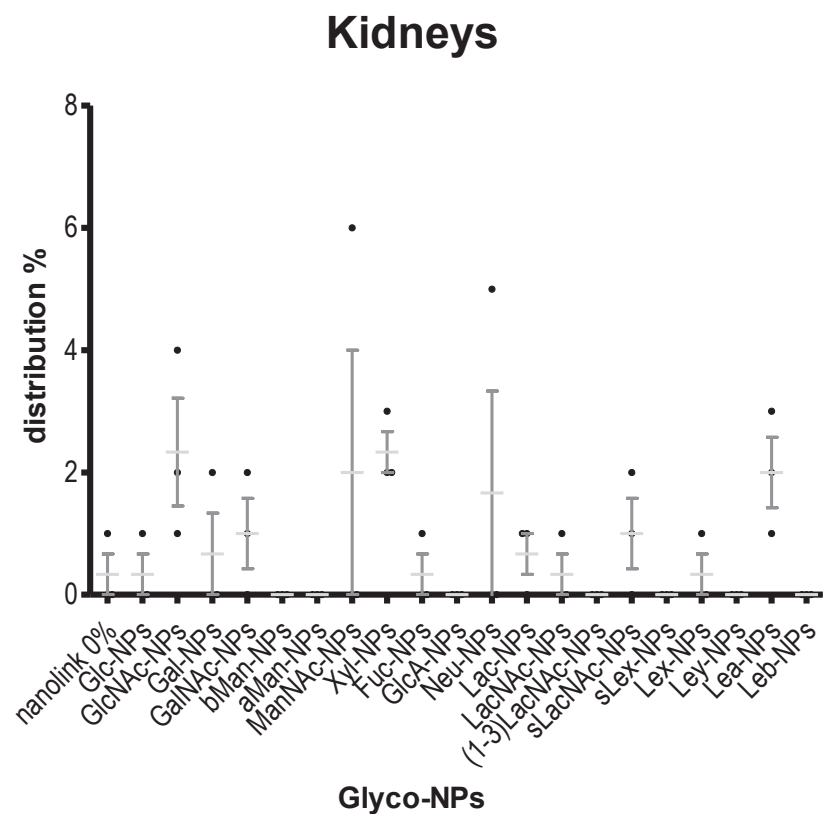
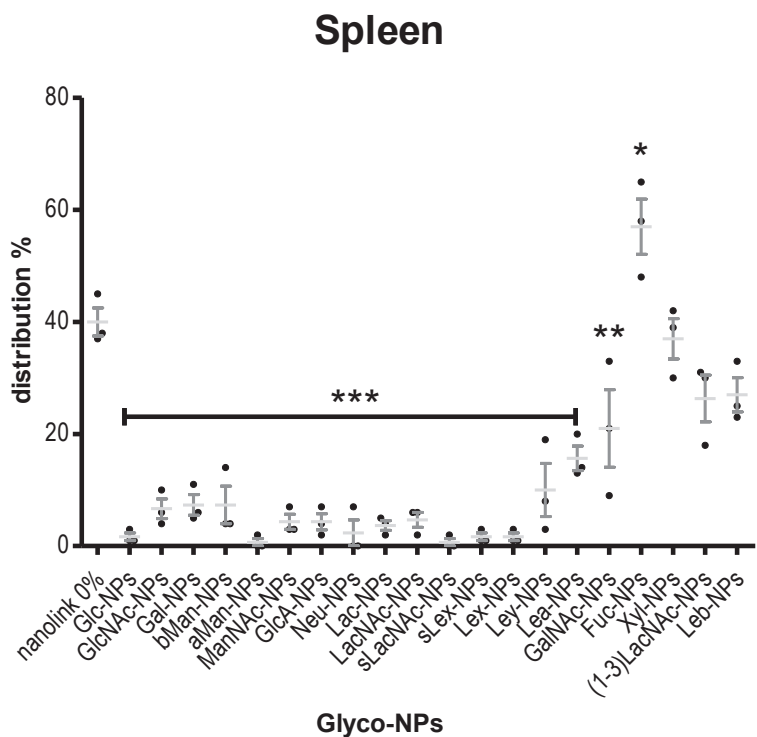


Figure S21 Comparison of the GNPs' distribution in the kidneys of CD-1 mice.



(Figure 6.6)

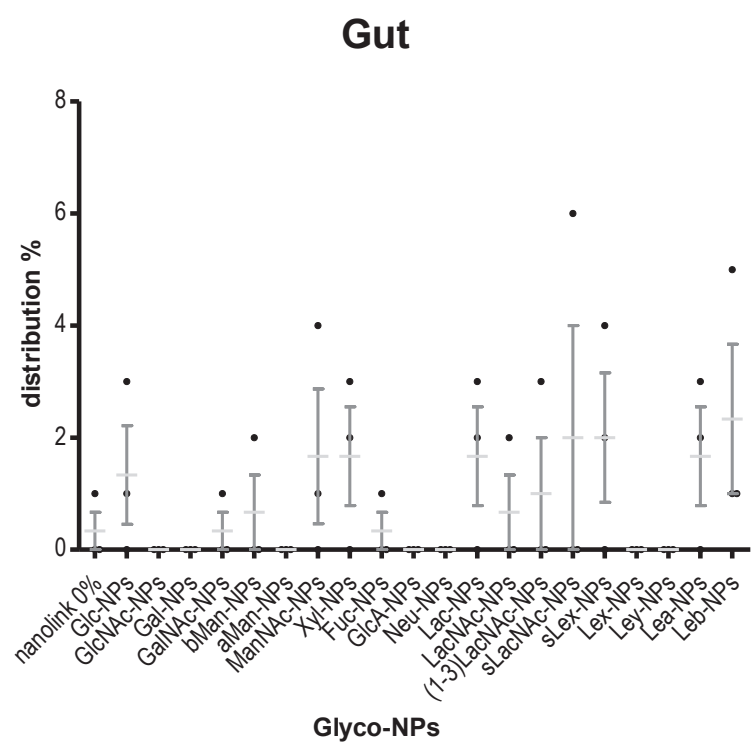
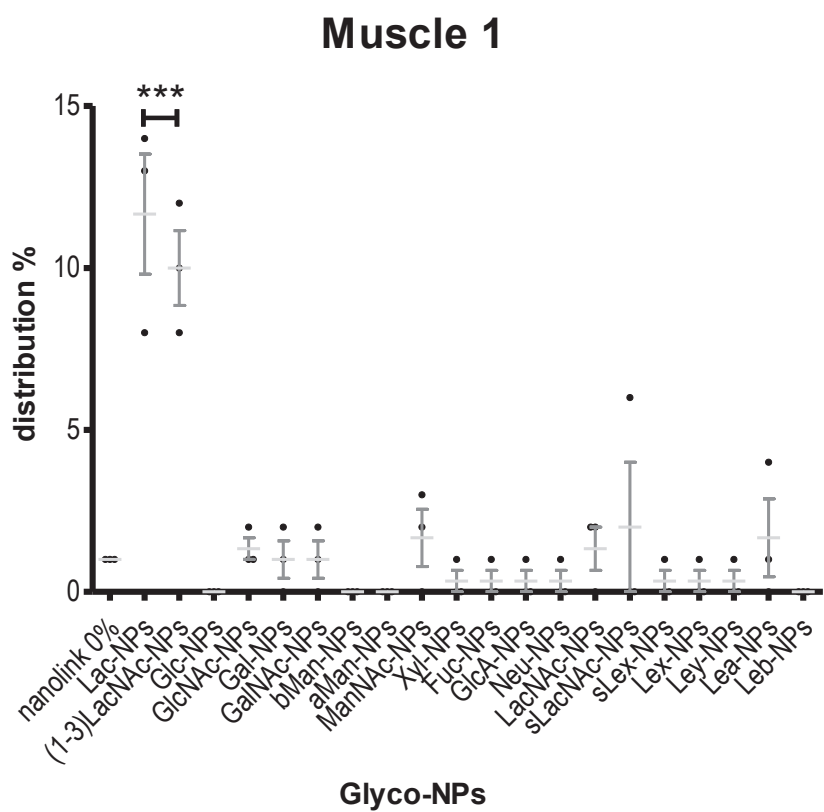


Figure S22 Comparison of the GNPs' distribution in the gut of CD-1 mice.



(Figure 6.9)



(Figure 6.8)

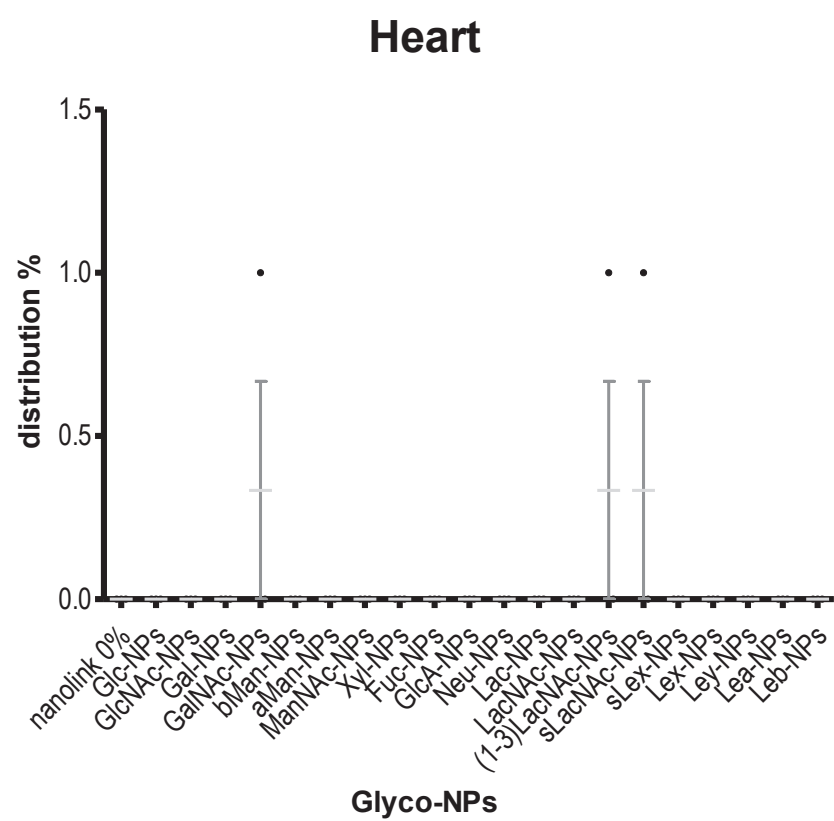
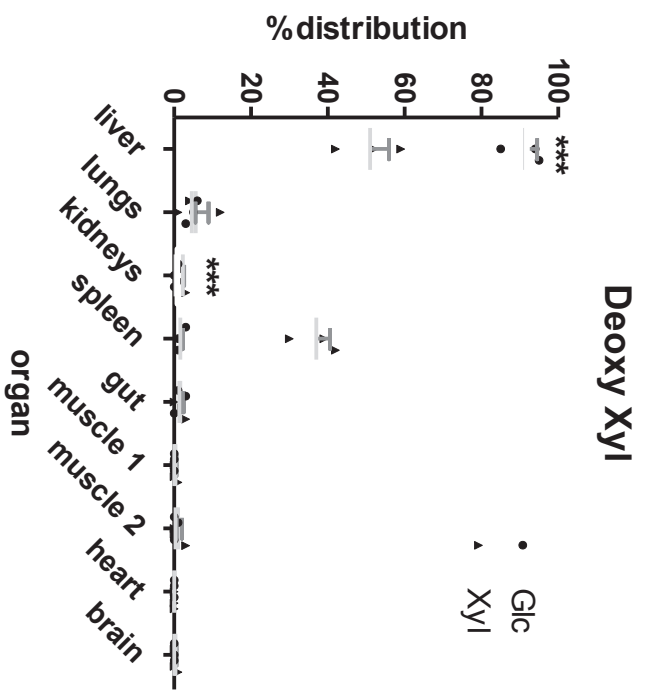
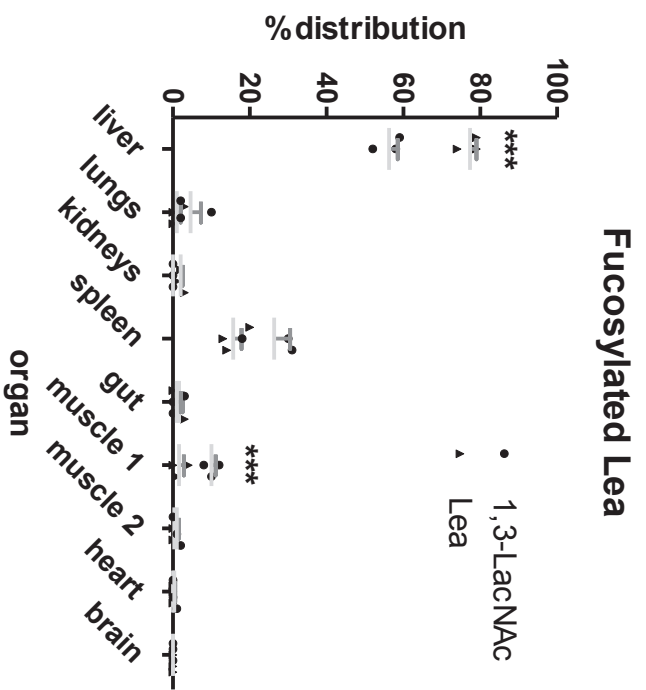


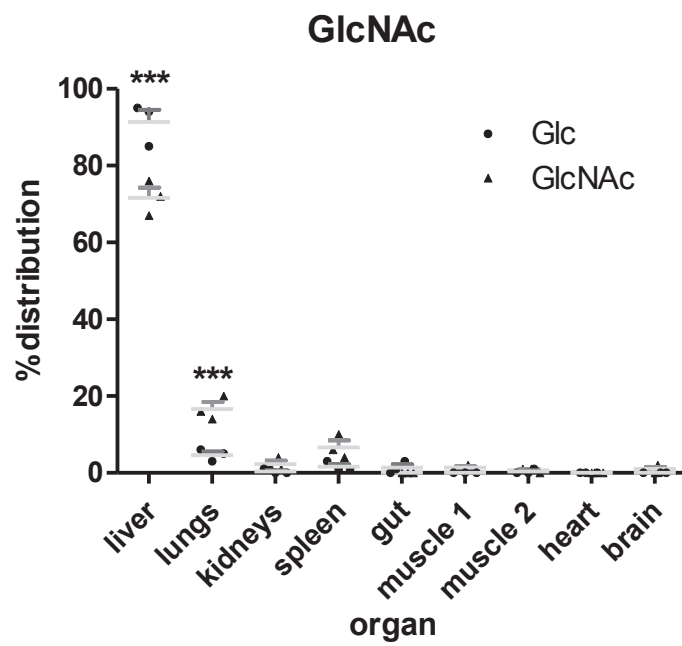
Figure S23 Comparison of the GNPs' distribution in the heart of CD-1 mice.



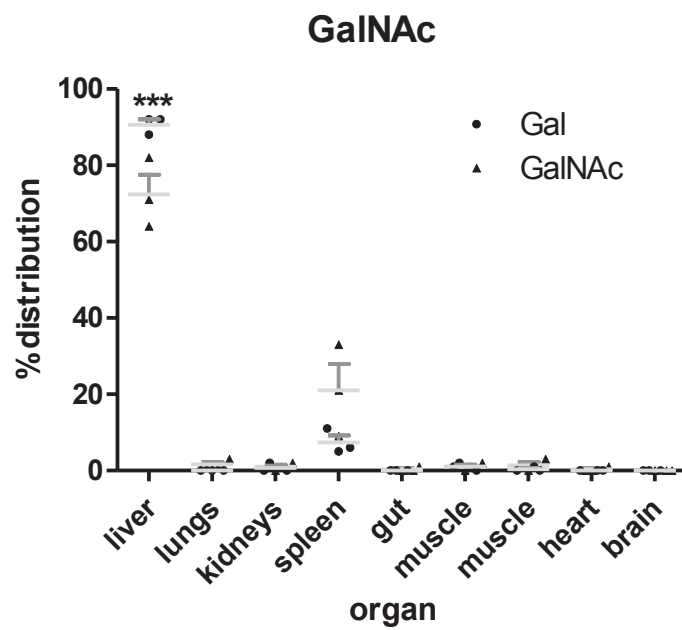
(Fig. 6.13)



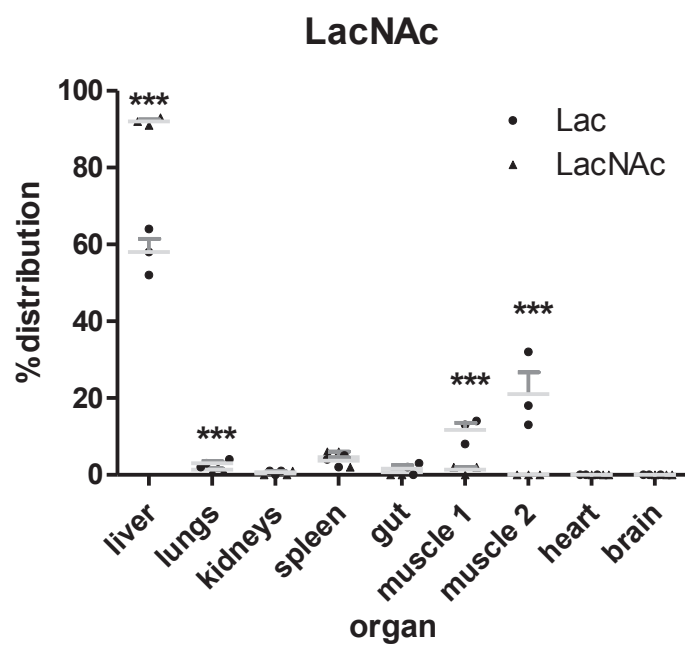
(Fig. 6.18)



(Fig. 6.16)



(Fig. 6.16)



(Fig. 6.16)

Histological studies

Sections of organs fixed in 4% PFA were embedded in wax, cut in 5 µm slices and deposited on slides. Slides were stained with nuclear fast red solution and Perl's staining, to detect the iron particles, using the following protocol:

1. Deparaffinise and hydrate sections

- 65 °C x 20 min
- xylene, 5 min x 2
- 100% EtOH, 5 min x 2
- 95% EtOH, 5 min
- 80% EtOH, 5 min
- distilled H₂O, 5 min

2. Stain

- mix 20% HCl with 10% K₄Fe(CN)₆ (1:1)
- stain for 20 min

3. Wash

- distilled H₂O, 5 min x 3

4. Counterstain

- nuclear fast red solution, 5 min

5. Rinse

- distilled H₂O, x 2

6. Dehydrate

- 95% EtOH, 5 min
- 100% EtOH, 5 min x 2

7. Clear

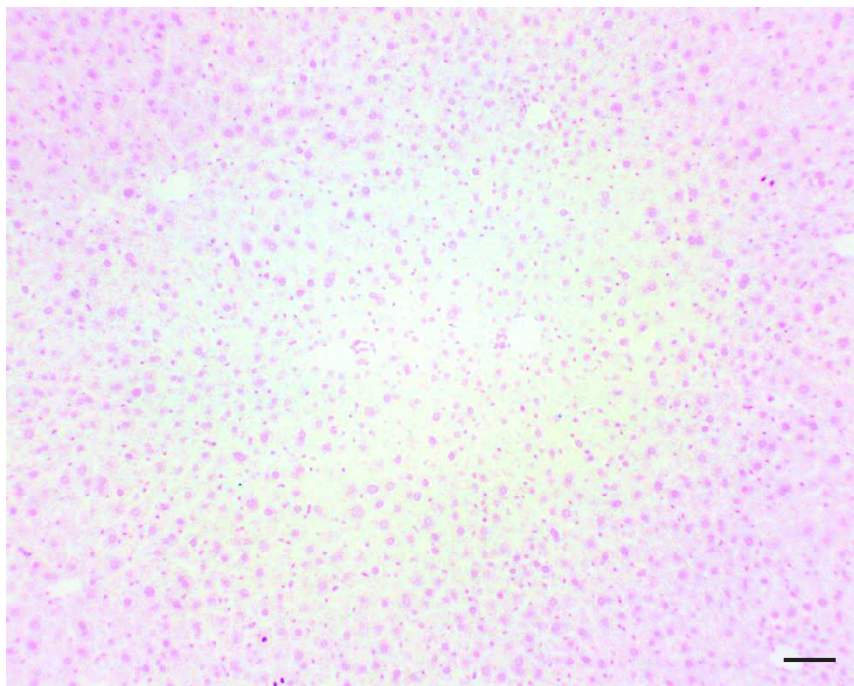
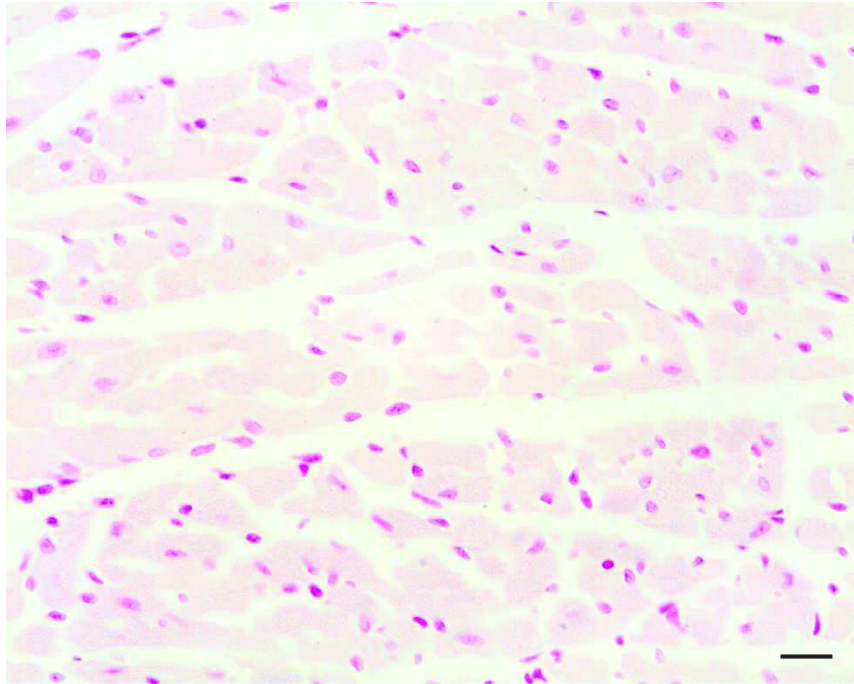
- xylene, 3 min x 2

8. Cover

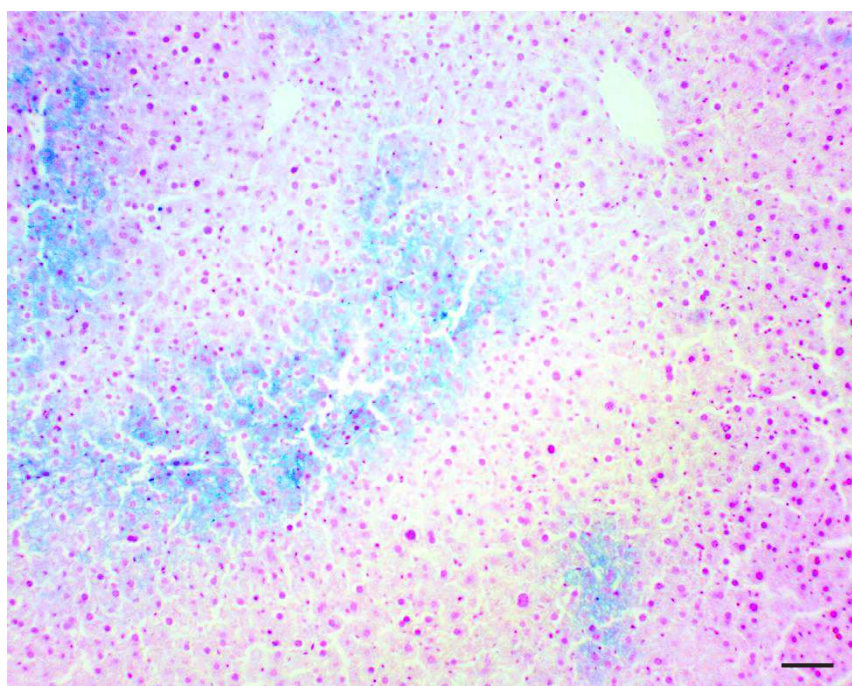
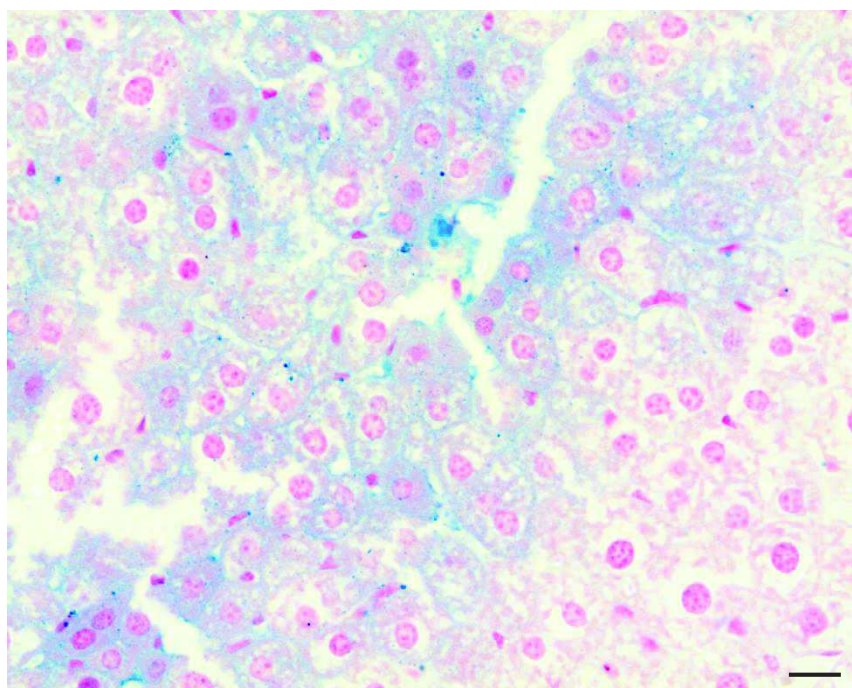
- coverslip with resinous mounting medium and allow to dry

Each organ has been shown as non-treated with GNPs and after treatment with GNPs (if GNPs have been detected after staining). For each sample two images were collected with different magnification: scale bars are 100 µm (upper images) and 50 µm (lower images), for all images.

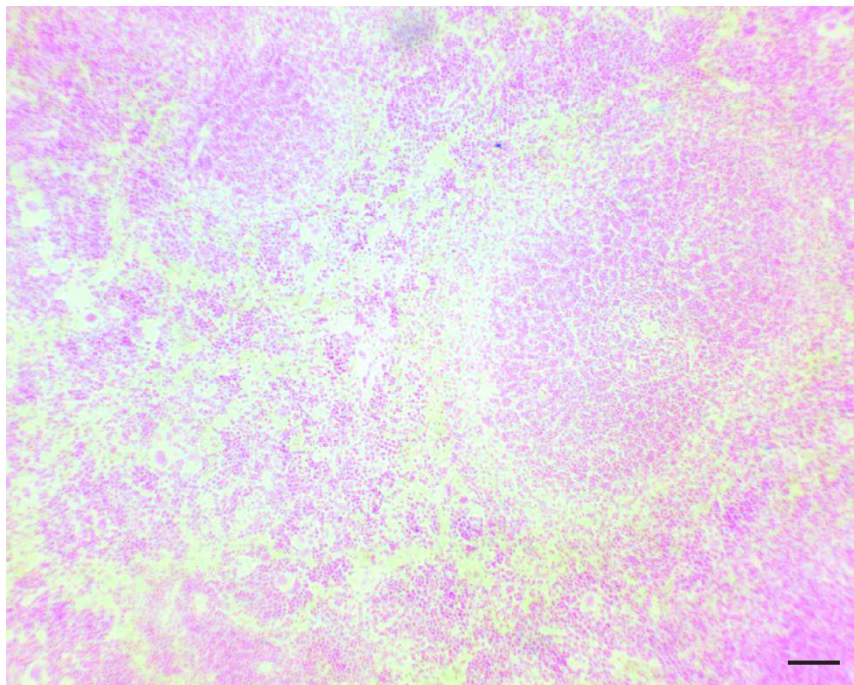
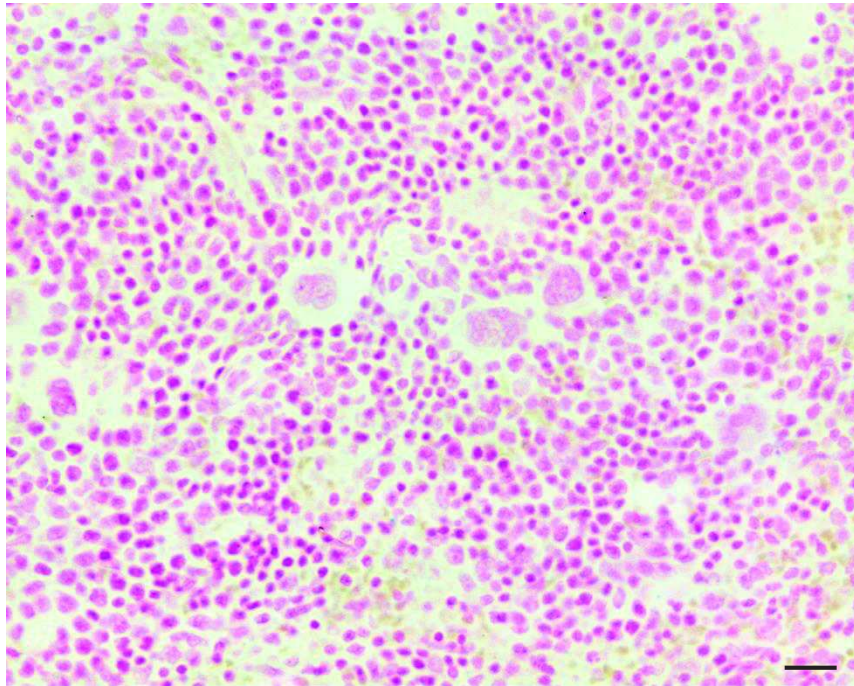
Liver



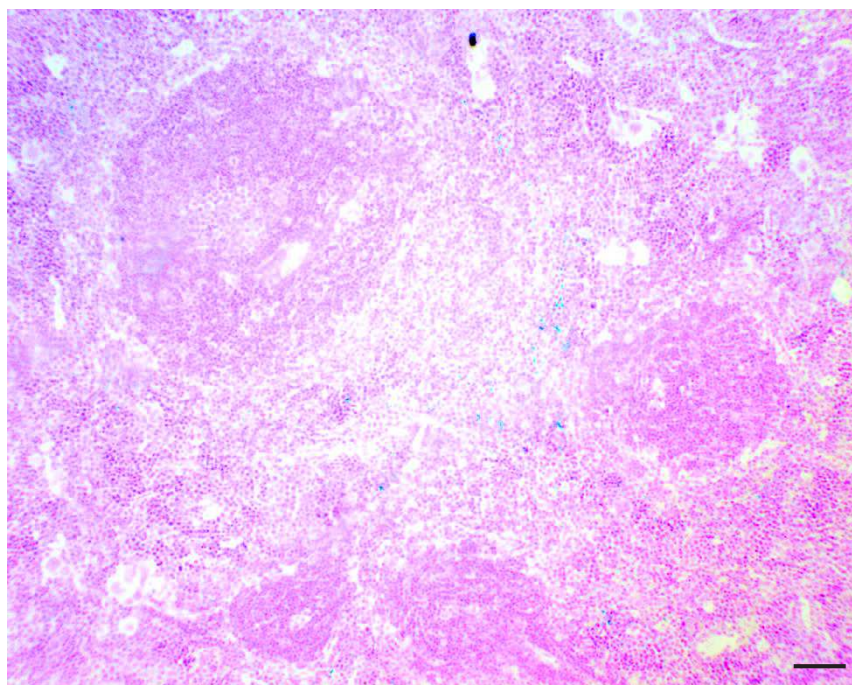
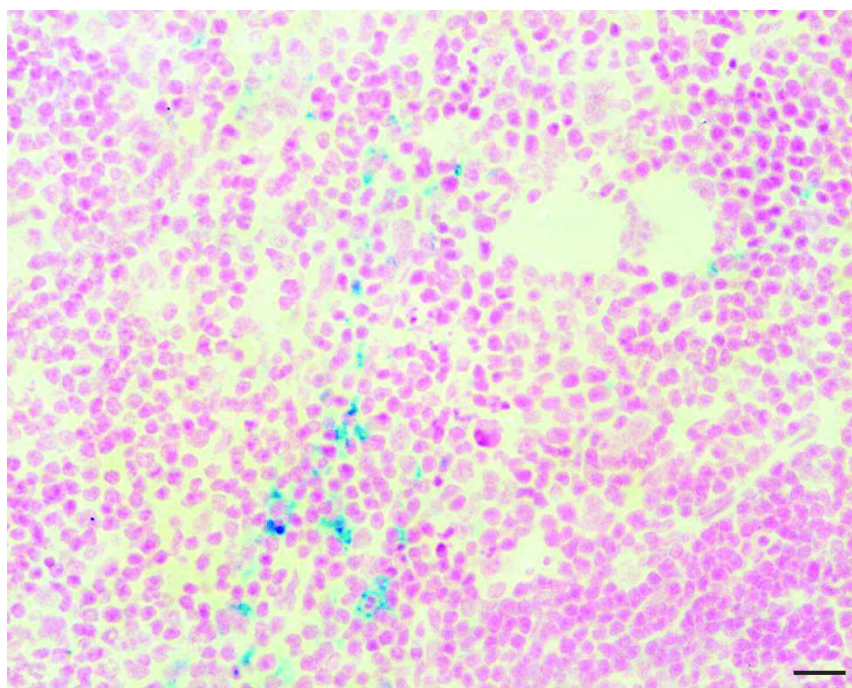
Liver + NanoLink



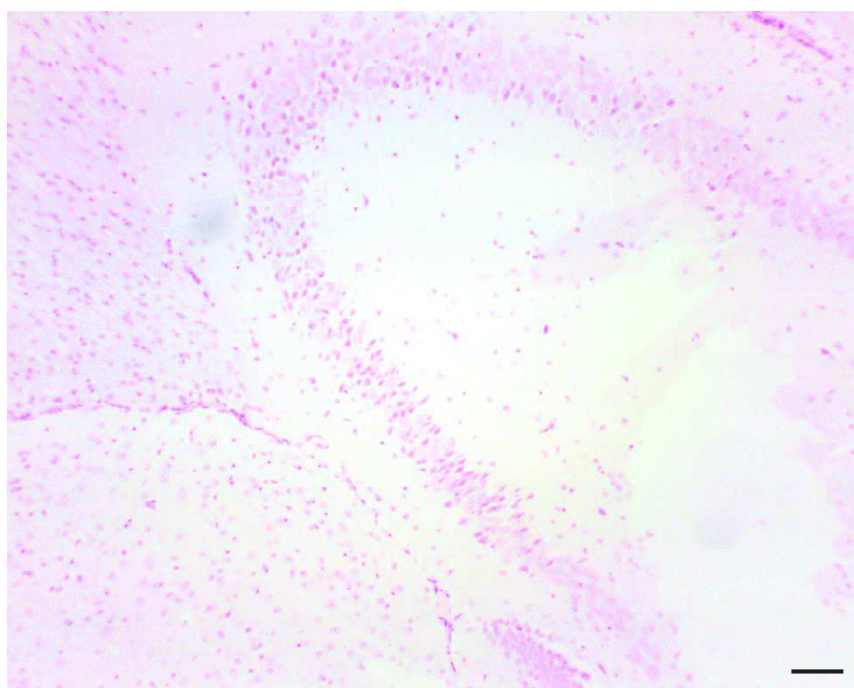
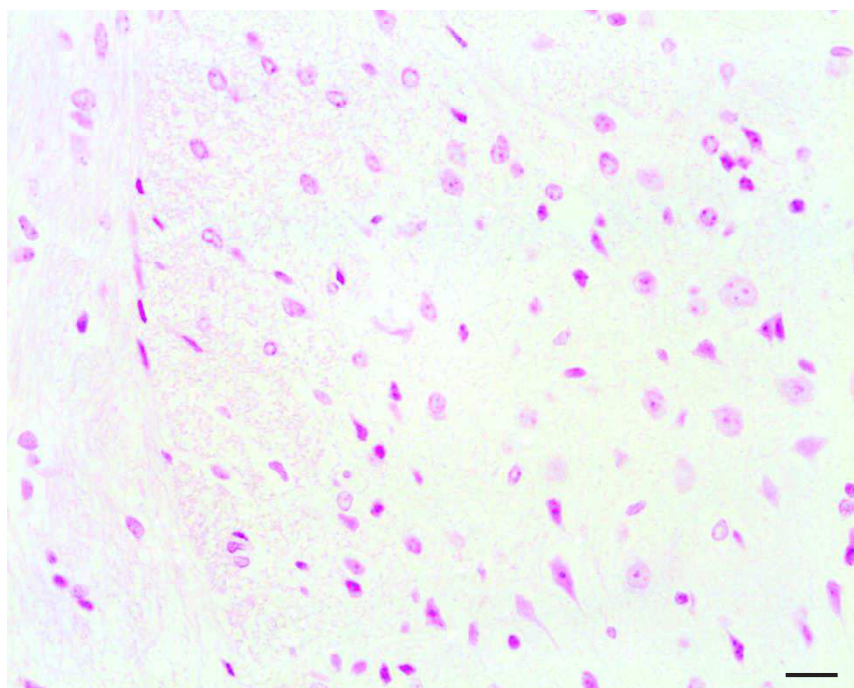
Spleen



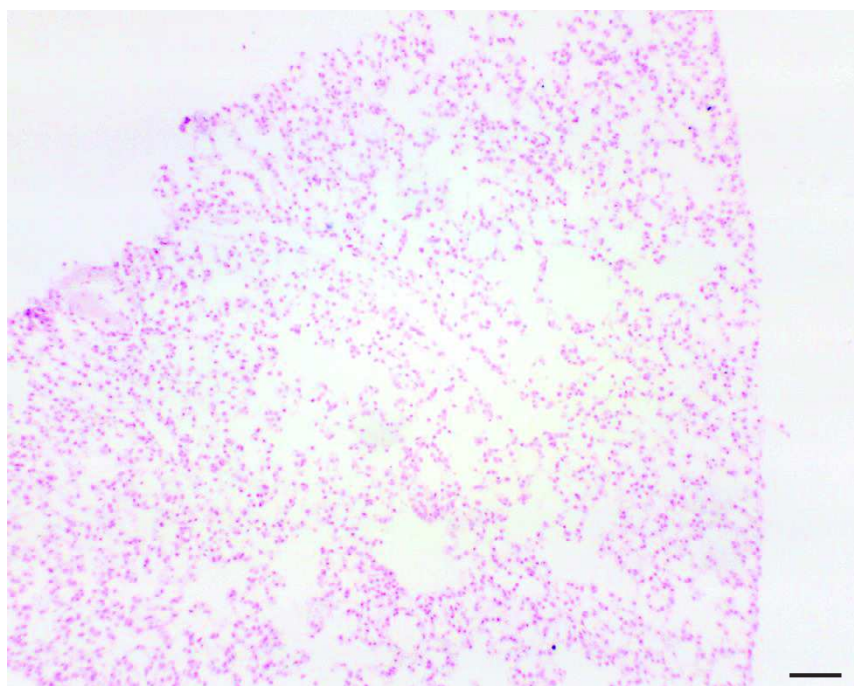
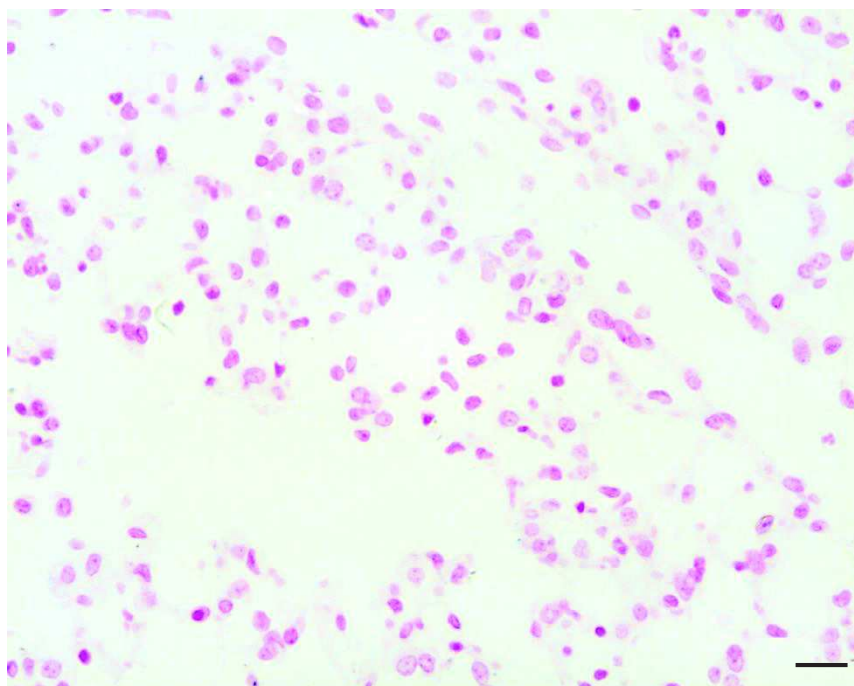
Spleen + Nanolink



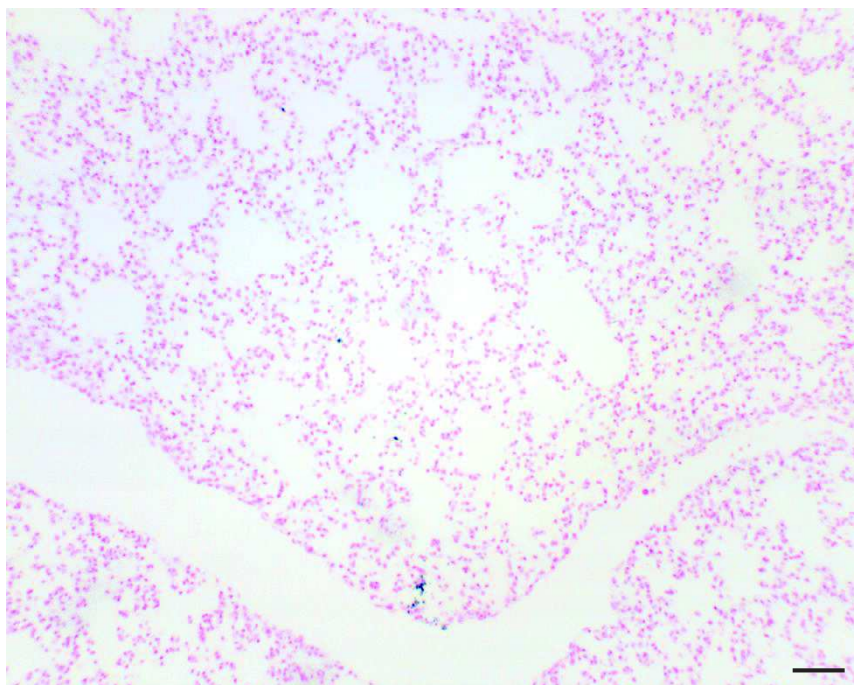
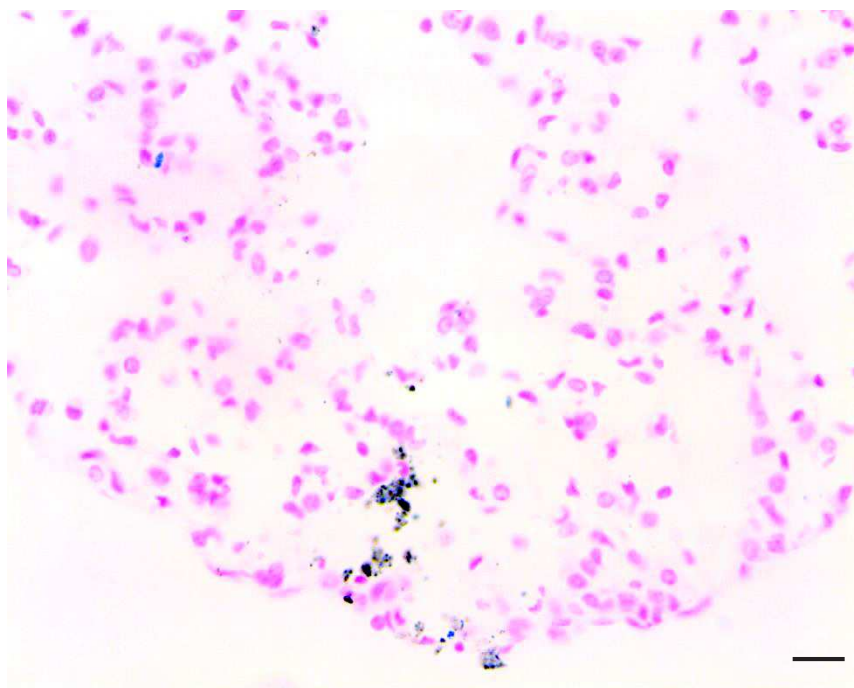
Brain



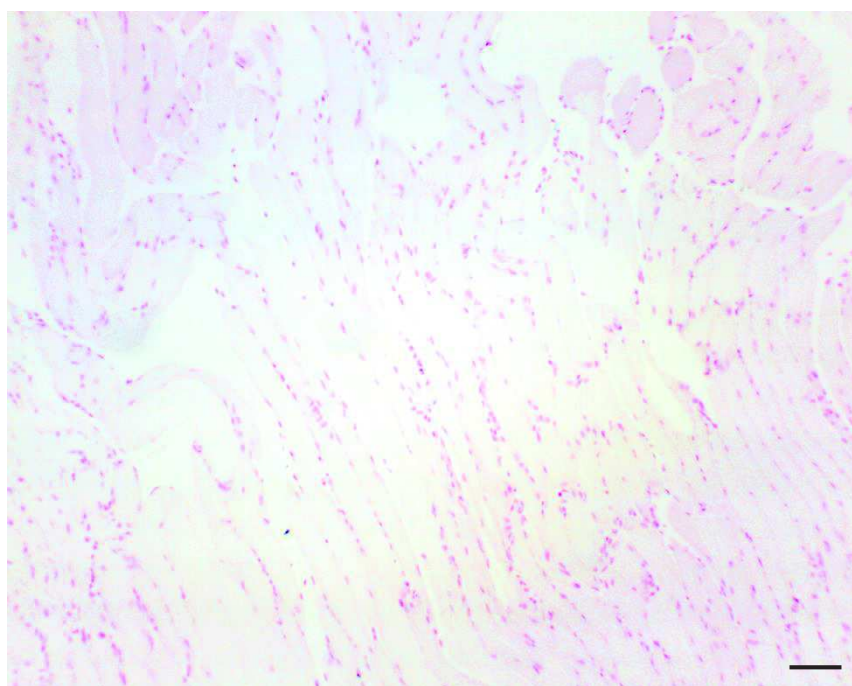
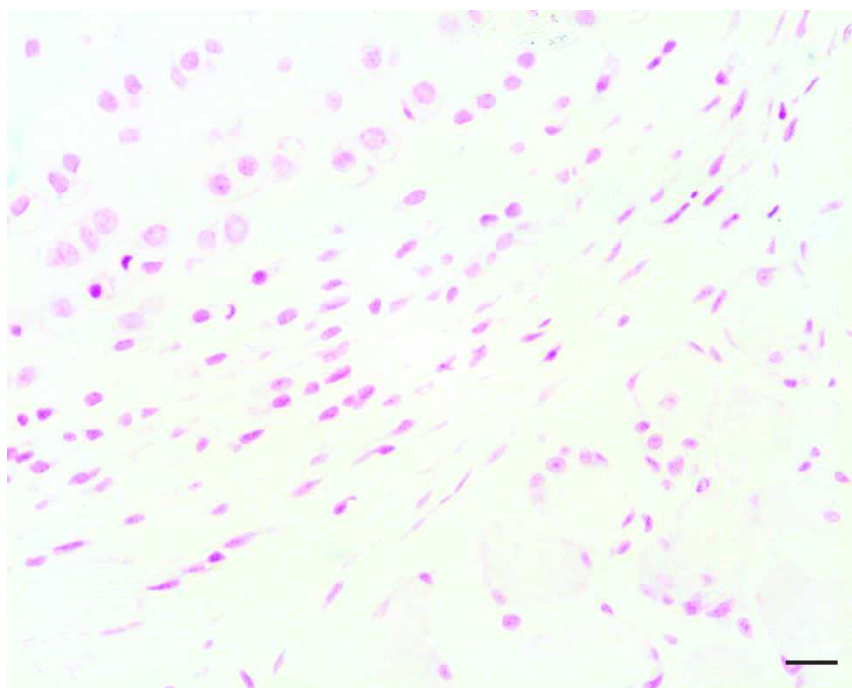
Lungs



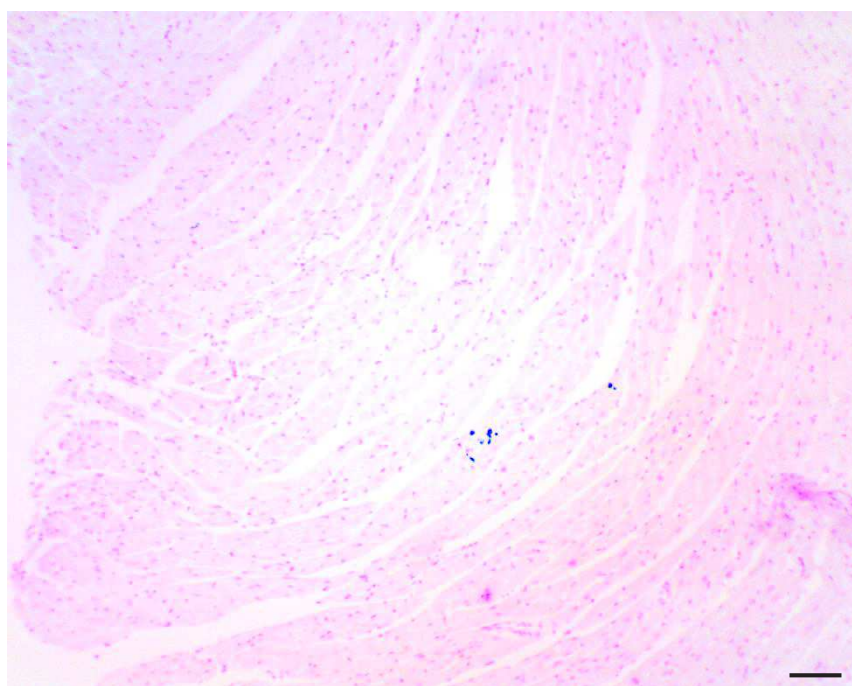
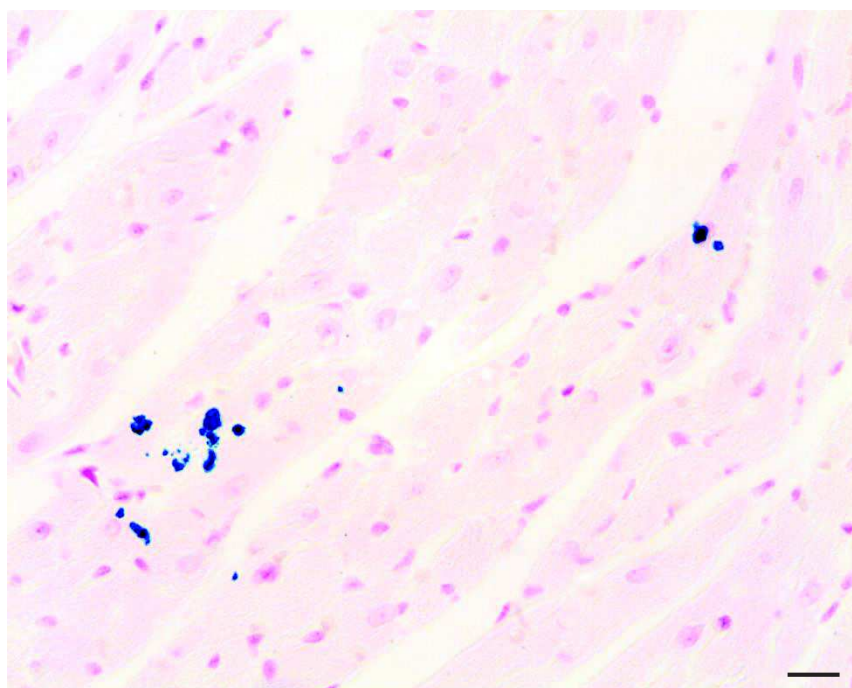
Lungs + GlcNAc



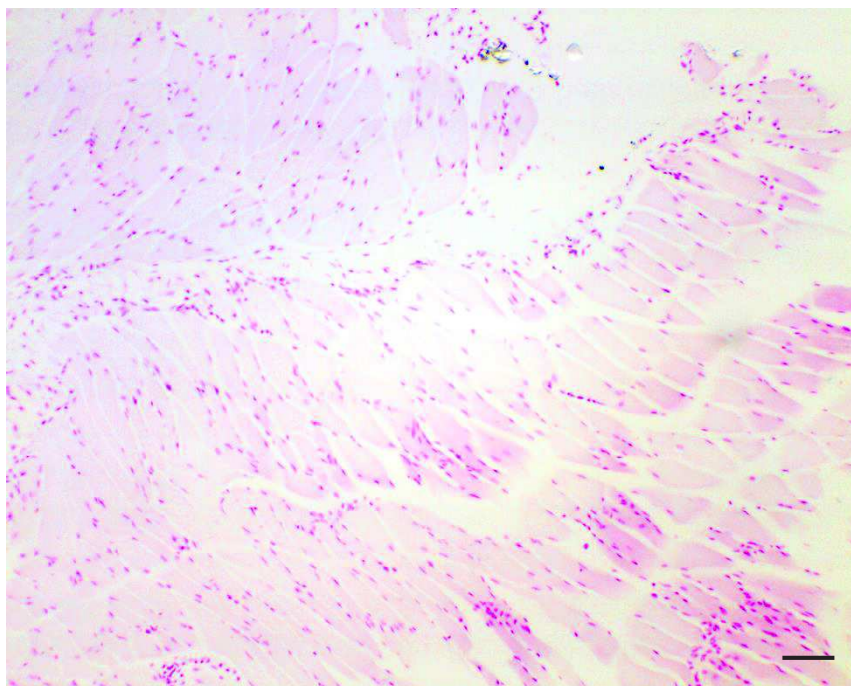
Muscle 1



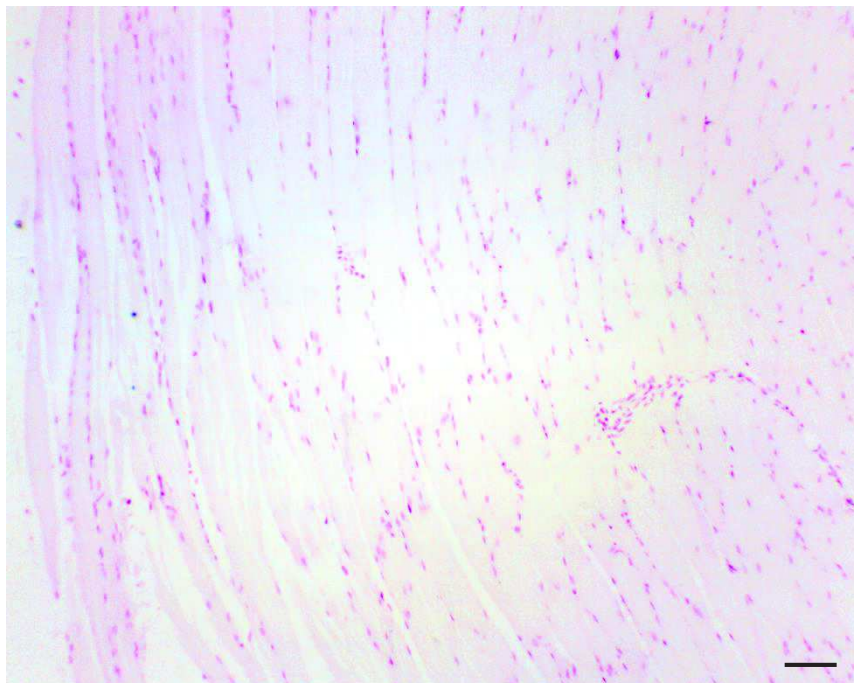
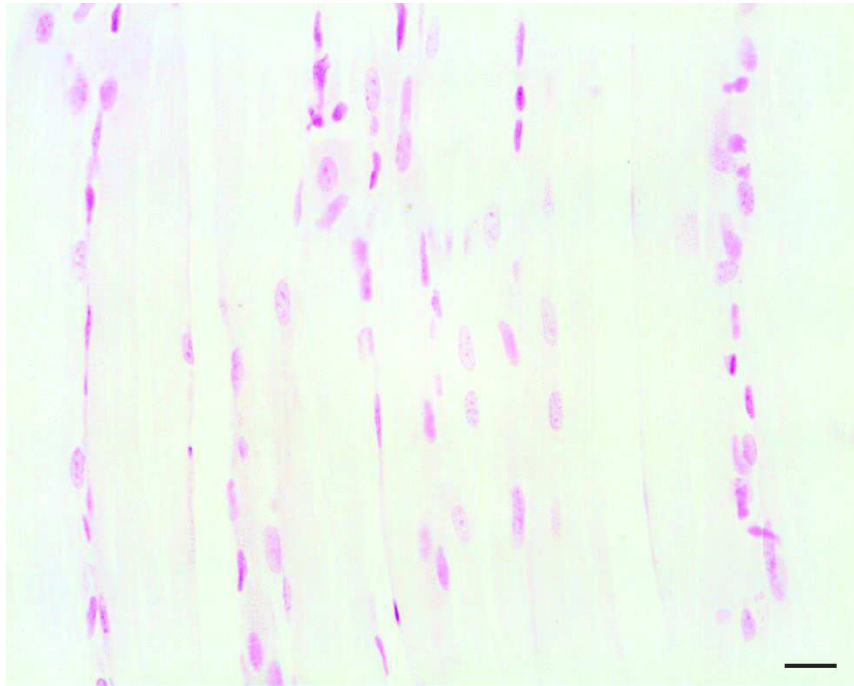
Muscle 1 + Lac



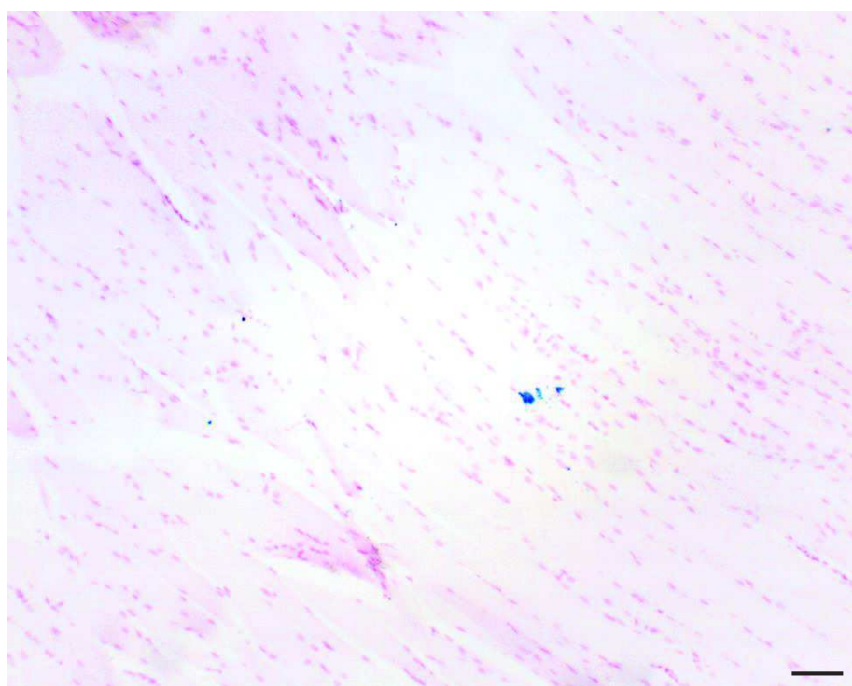
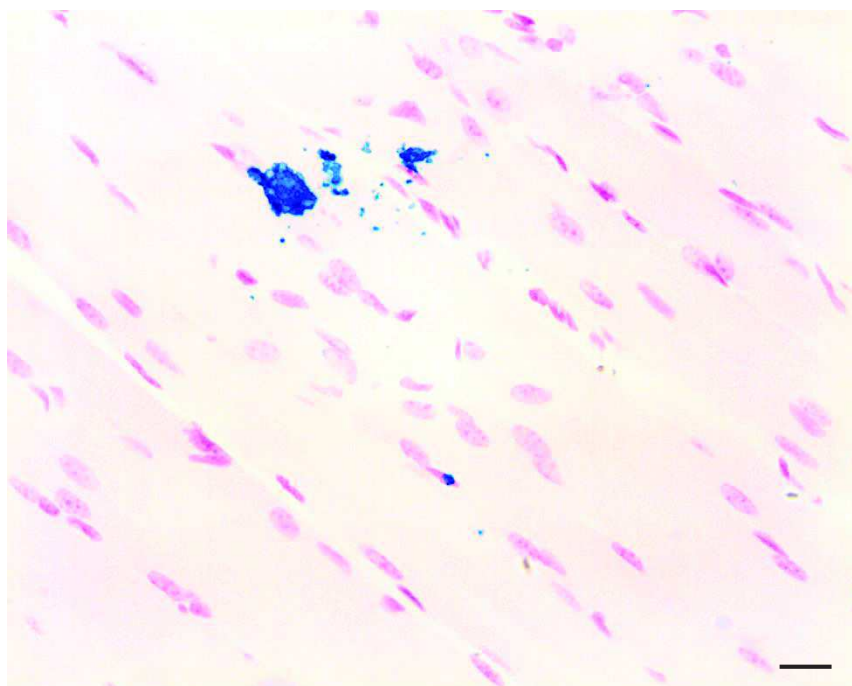
Muscle 1 + (1,4)-LacNAc



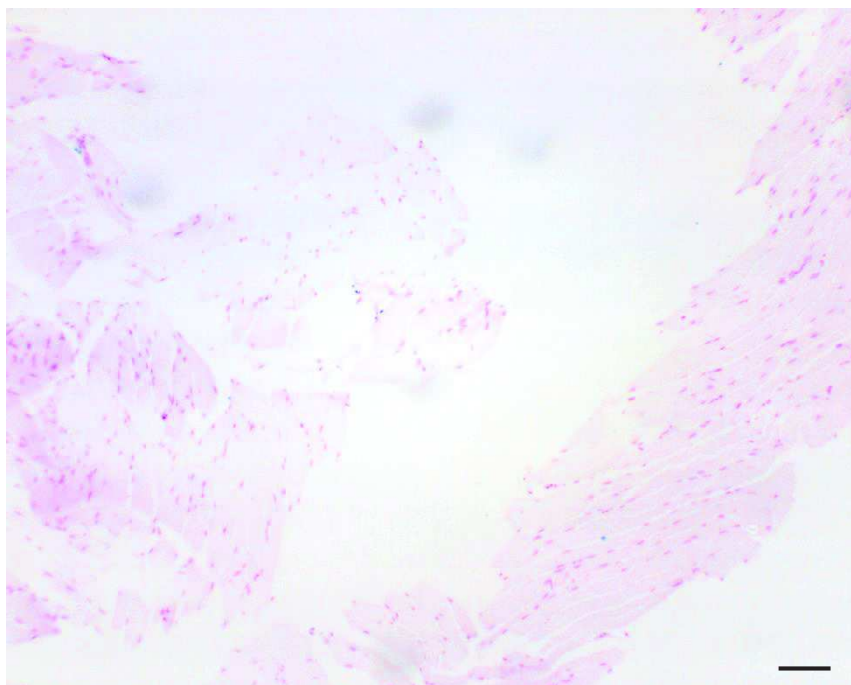
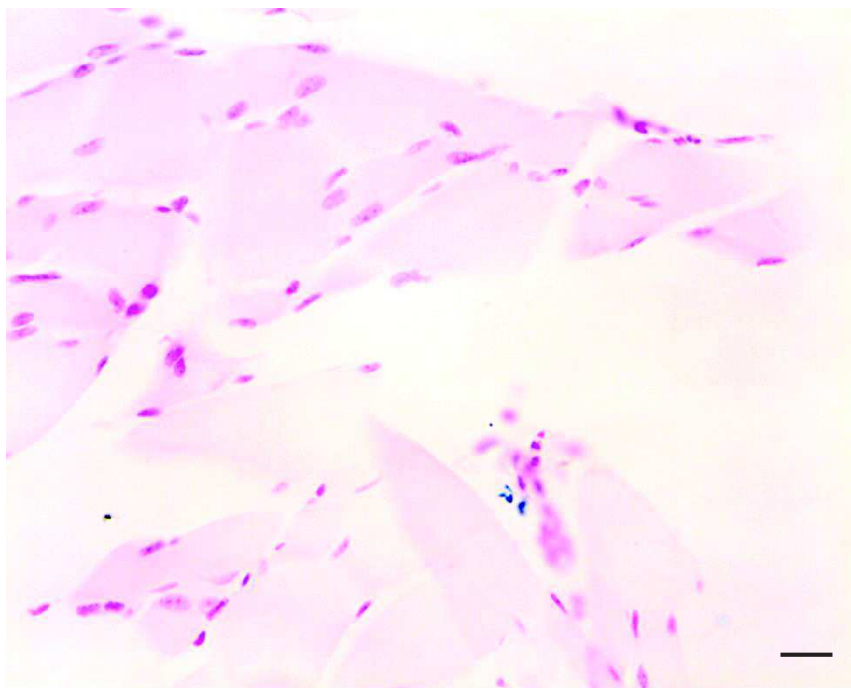
Muscle 2



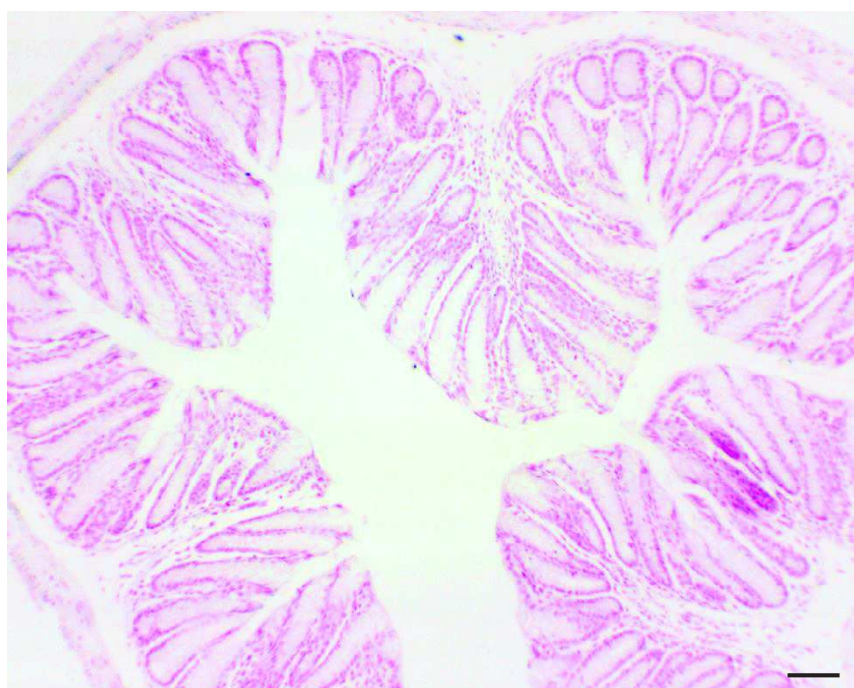
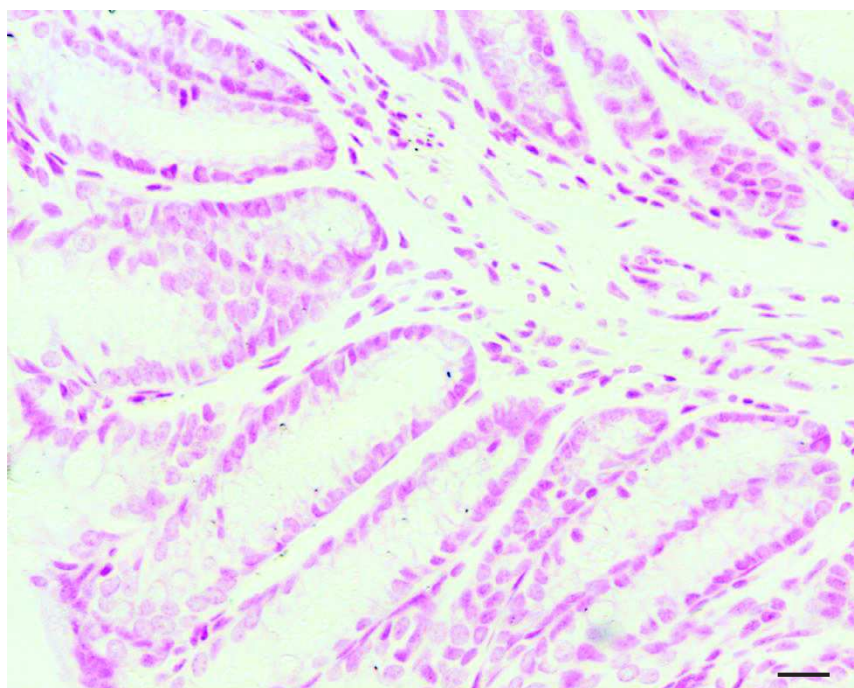
Muscle 2 + Lac



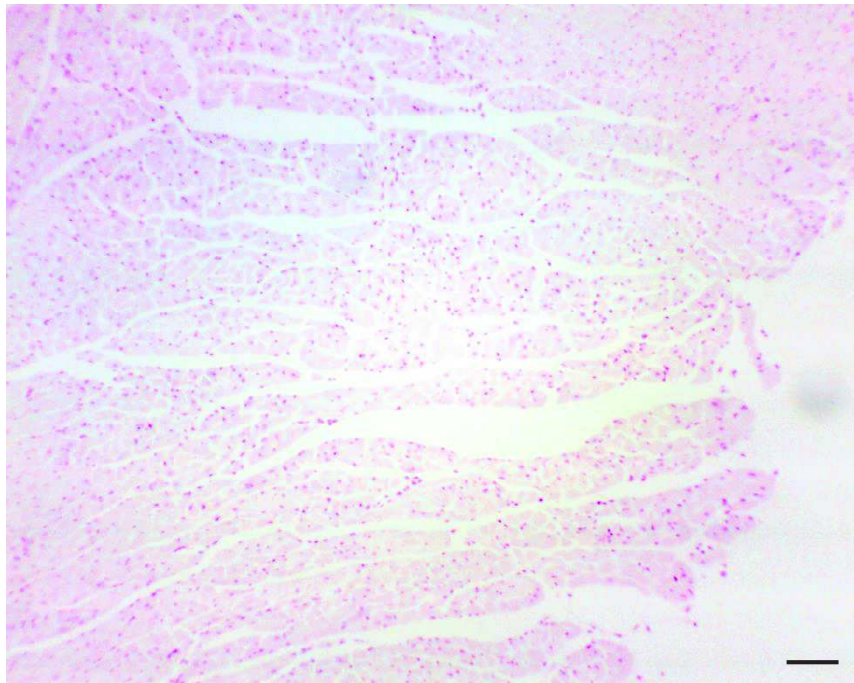
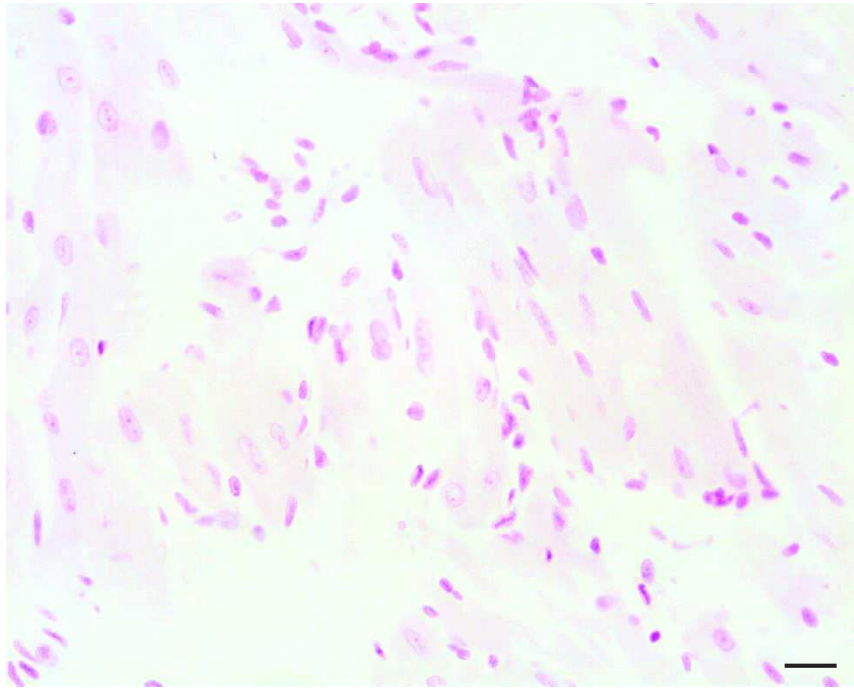
Muscle 2 + (1,4)-LacNAc



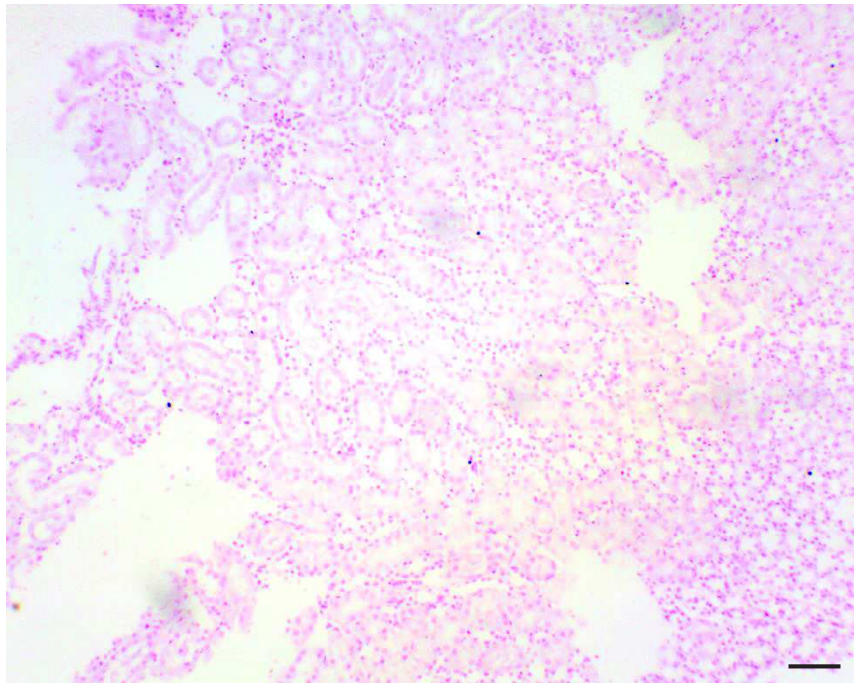
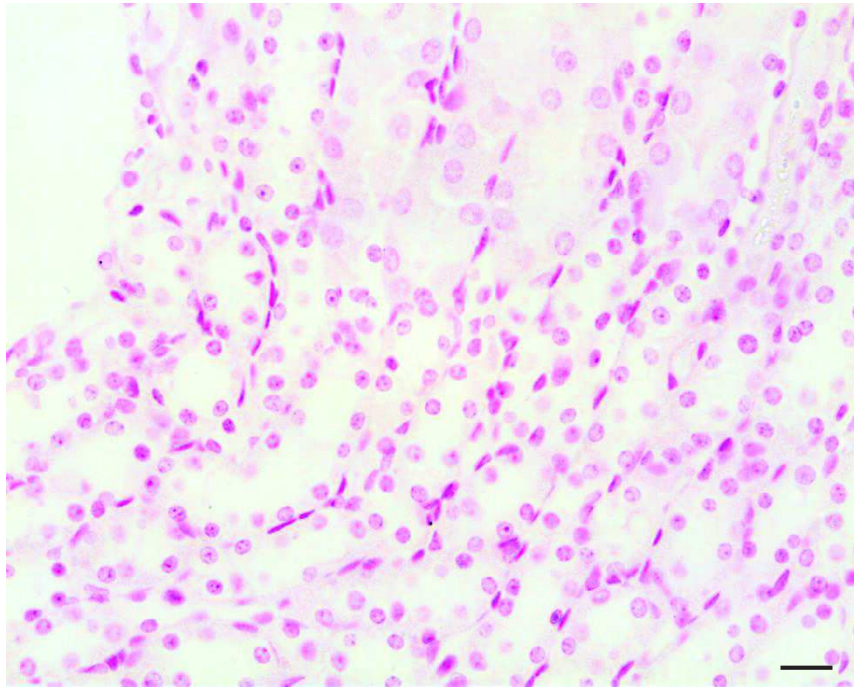
Gut



Heart



Kidneys



[¹⁸F]-NPs *in vivo* testing

Wistar rats (150 g body weight) were treated with IL-1 β (100 ng in 1 μ L of saline) by intracranial injection on the left hemisphere of the brain, under anaesthesia by isoflurane, and left to recover. After 4 h animals were injected *i.v.* with 0.2 mg of [¹⁸F]-NPs **4.24** and **4.25** (~4 mg nanoparticles/Kg rat) in 20 μ L of saline solution (0.9 % NaCl in H₂O), under anaesthesia by isoflurane. Animals were left to recover once again. After 30 min mice were killed by neck dislocation, the head removed and the brain harvested, frozen in isopentane/dry CO₂ bath and cut using cryostat in 25 μ m slices, subsequently exposed for autoradiography. All tests have been run in triplicate.

Na¹²⁵I@SWNTs *in vivo* testing

CD-1 male mice (20-25 g body weight) were injected *i.v.* with ~0.1 mg (~ 1 MBq) of Na¹²⁵I@SWNTs **5.46**, **5.47** and **5.48** in 100 μ L of saline solution (0.9 % NaCl in H₂O), under anaesthesia by isoflurane. The animals were left to recover. After 1 h mice were killed by CO₂ asphyxiation. Sections of liver (Li), lungs (Lu), heart (H), spleen (S), kidneys (K), gut (G), muscles (M1 from hind leg and M2 from front leg) and brain (B) were harvested for *ex-vivo* studies of the distribution of the nanotubes. All tests have been run in triplicate.

mouse				sample					
				1	Na ¹²⁵ I@SWNTs				
				2	Na ¹²⁵ I@SWNTs				
				3	Na ¹²⁵ I@SWNTs				
				4	GlcNac-Na ¹²⁵ I@SWNTs				
				5	GlcNac-Na ¹²⁵ I@SWNTs				
				6	GlcNac-Na ¹²⁵ I@SWNTs				
				7	Gal-Na ¹²⁵ I@SWNTs				
				8	Gal-Na ¹²⁵ I@SWNTs				
				9	Gal-Na ¹²⁵ I@SWNTs				
mass (mg)									
liver	lungs	kidneys	spleen	gut	muscle 1	muscle 2	heart	brain	
64.3	52.3	49.6	40.3	66.1	67.4	67.2	59.3	65.9	
59.0	55.6	50.1	47.4	62.3	66.5	66.5	58.2	80.3	
65.6	49.8	51.9	49.1	59.8	67.1	65.4	57.7	75.3	
62.8	56.4	49.3	45.8	60.0	59.9	66.7	56.8	73.7	
59.9	54.7	48.7	44.7	59.9	64.3	68.1	58.3	74.1	
60.6	52.0	50.0	48.2	59.3	62.6	59.8	60.0	69.8	
52.5	52.1	49.8	50.0	60.3	67.5	68.8	57.3	71.2	
68.4	55.4	52.1	40.9	61.0	66.6	62.6	55.7	70.9	
65.7	49.0	50.5	48.9	61.2	69.1	63.1	53.5	78.7	
counts per minute (CPM)									
234	6871	109	249	139	129	139	118	228	
335	7691	238	328	197	237	118	98	97	
125	8320	117	217	187	166	189	148	129	
117	8593	127	180	218	257	210	217	100	
235	8344	216	197	209	248	276	100	210	
436	7686	226	114	215	229	255	184	164	
122	7762	122	132	223	170	157	175	143	
234	8237	243	173	194	121	308	96	241	
328	9202	221	215	186	322	185	216	221	

CPM/g tissue								
0.00	0.13	0.00	0.01	0.00	0.00	0.00	0.00	0.00
0.01	0.14	0.00	0.01	0.00	0.00	0.00	0.00	0.00
0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.01	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00

% ID/g								
2	85	1	4	1	1	1	1	2
3	83	3	4	2	2	1	1	1
1	89	1	2	2	1	2	1	1
1	86	1	2	2	2	2	2	1
2	84	2	2	2	2	2	1	2
4	83	3	1	2	2	2	2	1
1	88	1	2	2	1	1	2	1
2	84	3	2	2	1	3	1	2
2	86	2	2	1	2	1	2	1

organ	%ID/g		
	Na ¹²⁵ I@SWNTs	GlcNac-Na ¹²⁵ I@SWNTs	Gal-Na ¹²⁵ I@SWNTs
liver	2	2	2
lungs	85	84	86
kidneys	2	2	2
spleen	3	2	2
gut	2	2	2
muscle 1	2	2	2
muscle 2	1	2	2
heart	1	2	2
brain	1	1	1

Fullerenes as support for oligosaccharide synthesis

7.1 Introduction

7.1.1 Solid-phase synthesis of oligosaccharides

In the past few decades, the introduction of solid-phase peptide synthesis (SPPS)^[1] represented a ground-breaking event in peptide chemistry. The possibility of synthesising complex peptides with such a simple and versatile technique opened the possibility to study their potential in many fields. Other than earning to its discoverer the Nobel Prize in 1984, the importance of this invention is underlined by the incredible number of discoveries and publications that came in the years that followed. However, tackling the multistep synthesis of other complex molecules is nowadays still a goal to look forward to. A striking example is the synthesis of oligosaccharides. In oligosaccharide (OS) synthesis a sugar glycosylating reagent (donor) reacts with a nucleophile (acceptor). The acceptor is usually a protected sugar bearing a free OH group in one (or more) positions. One of the reasons for the complexity of OS synthesis is that the chemistry of protecting groups (PGs) has to be finely tuned to allow the site selective modification of the acceptor and selective deprotection of the oligosaccharide obtained to further react with another donor and elongate the carbohydrate chain.

Moreover, glycosylation has to be stereoselective so it will not yield a mixture of different and often inseparable isomers. The solution-phase synthesis of OS is a lengthy and expensive process due to the multiple deprotecting and coupling steps, each followed by extensive and challenging purifications. Solid-phase synthesis of oligosaccharides (SPOS) is desirable as it would minimise and simplify the frustrating and time consuming purification steps. This methodology requires an insoluble tag (or support) functionalised with a linker that needs to be chemically inert during the synthesis but can be readily cleaved from the resin at the end of the reaction sequence. The tag can support either the donor^[2-4] or the acceptor^[5] or both. A first approach toward the SPOS was presented by Seeberger and co-workers in 2001^[5] (Fig. 7.1).

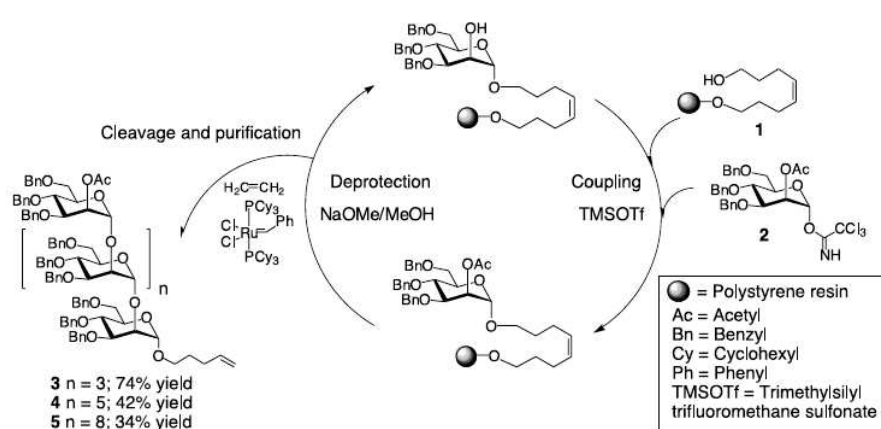


Figure 7.1 Polystyrene resin supported solid-phase synthesis of OS.^[5]

With his automated synthesis, Seeberger managed to synthesise β -phytoalexin elicitor in 10 h and 80% yield, compared to the synthesis published by Nicolau^[6] whose efforts gave the same compound in 122 h with only 20% yield. Although this method represented the most advanced technique explored so far, it is just a first step on the road to achieving an efficient and reliable strategy for sugar

synthesis. Amongst the drawbacks that still need to be addressed are the excess of glycosyl donor required, the poor reactivity at the solid-solution interface,^[7] and the difficulty in tracking the efficiency of each glycosylation.

7.1.2 Fullerenes as support for SPOS

In our group, fullerenes are under investigation as new phase tagging agents for glycoside synthesis. This system consists of C₆₀ fullerenes functionalised with a SH-terminating linker (C₆₀-SH) on which carbohydrates can be added. This linker should not be affected during the synthesis, does not interfere with the glycosylation/deprotection steps (elongation of the carbohydrate chain) and can be promptly cleaved at the end of the synthesis, releasing the final compound (Fig. 7.2).

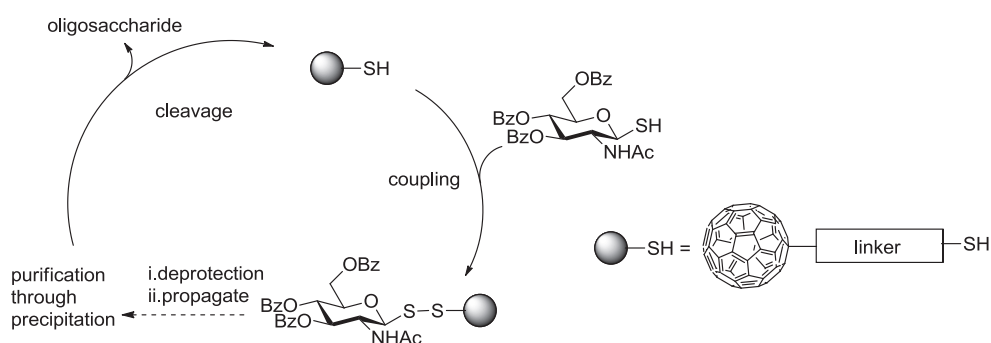
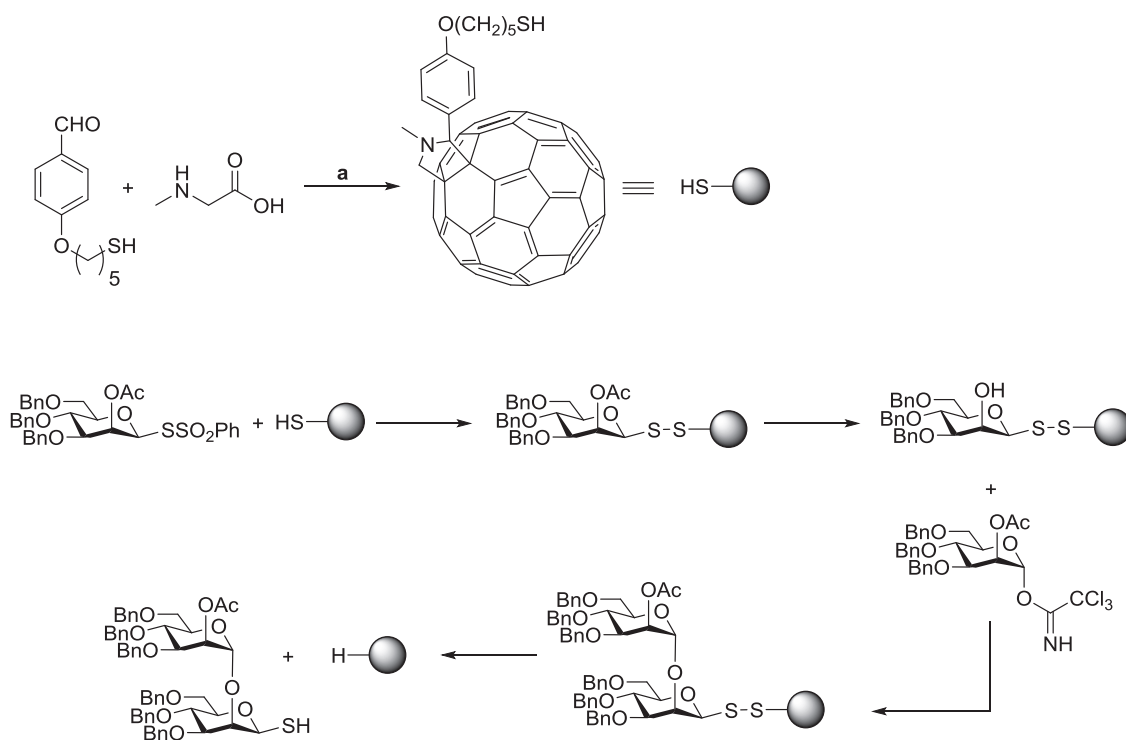


Figure 7.2 Fullerene supported solid-phase synthesis of OS.

The advantage of using C₆₀ as support is the ability to perform homogenous solution-phase synthesis with all the advantages of a solid-phase work-up through precipitation. Also, fullerene derivatives can be characterised by

conventional organic microanalysis (NMR, mass), and the synthesis can be easily followed step-by-step on TLC without specific visualisation methods.

An early attempt was carried out synthesising a fulleropyrrolidine derivative (Scheme 7.1).



Scheme 7.1 Fulleropyrrolidine tagging agent and plan for the solid-phase synthesis of thio-mannobiose: a) C₆₀, toluene, 130 °C, 15%.

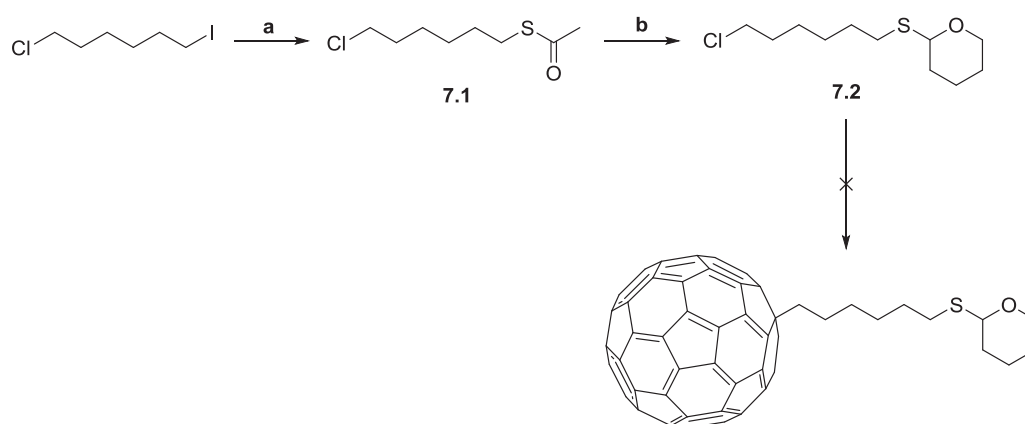
The fulleropyrrolidine derivative was successfully synthesised, using the Prato reaction^[8] to functionalise the fullerene. The plan was to couple the mannosyl-phenylthiosulphonate acceptor on the fullerene and, after deprotection, perform the first glycosylation with mannosyl trichloroacetimidate (donor). The di-mannose derivative obtained would have been cleaved from the carbon tag with tributylphosphine, regenerating the free thiol on the fullerene.

Unfortunately, this system proved not to be well suited for various reasons: glycosylations had to be carried out in acidic conditions, as the disulphide linkage is susceptible to bases; the tertiary amine of the fulleropyrrolidine is a potential site to react with Lewis acids used during the glycosylation; the yields obtained were not competitive with the methods described by Seeberger.^[5]

7.2 Synthesis of fullerene tags

7.2.1 First approach: functionalisation with a Grignard reagent

The idea of using fullerenes as support for SPOS seemed promising although a more robust approach to install the spacer on the fullerenes was needed. Our first attempt to synthesise thiol-terminated fullerenes without the pyrrolidine ring was inspired by the procedure designed by Kang et al. in 1998.^[9] This approach consisted of reacting a Grignard reagent, specifically a tetrahydropyran-2-ylsulfanylalkyl magnesium chloride, with C₆₀ to make mono-substituted fullerenes (Scheme 7.2).

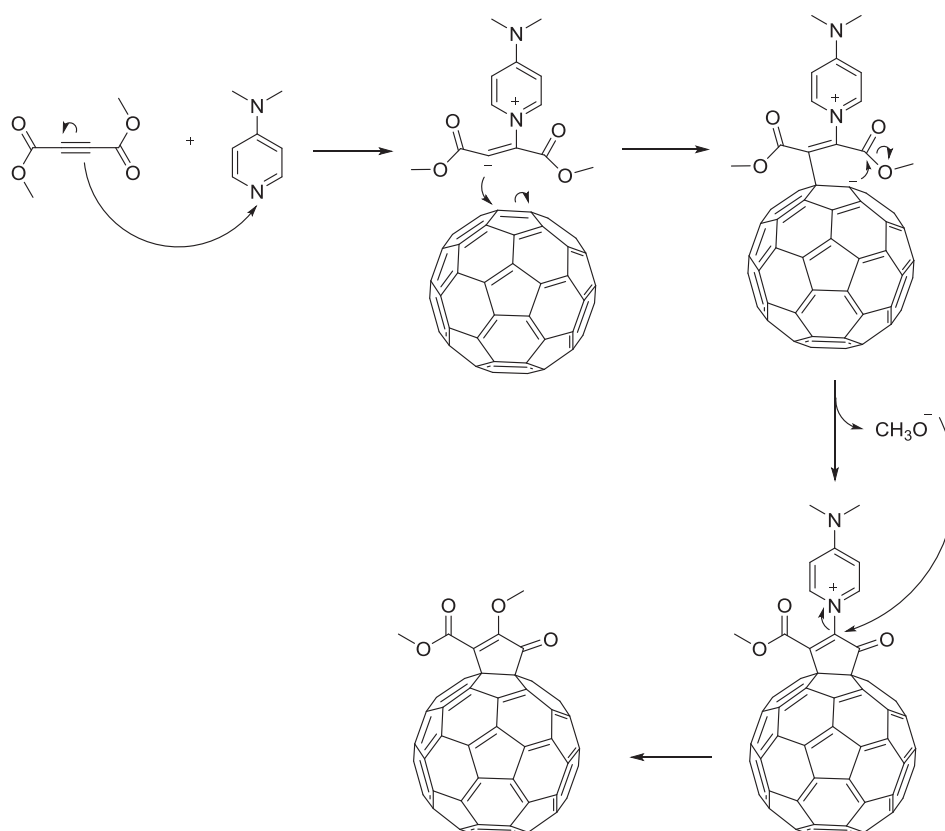


Scheme 7.2 Reagents and conditions: a) KSAc, acetone, r.t., 98%; b) BF₃Et₂O, DHP, MeOH, r.t., 40%.

Even though we synthesised the linker, grafting it onto the fullerenes was unsuccessful. We transformed the halogenated starting material to the organomagnesium derivative and added fullerenes *in situ* preserving strict dry conditions but no product was observed. Different equivalents, solvents and temperatures were used but still we could not obtain the desired compound.

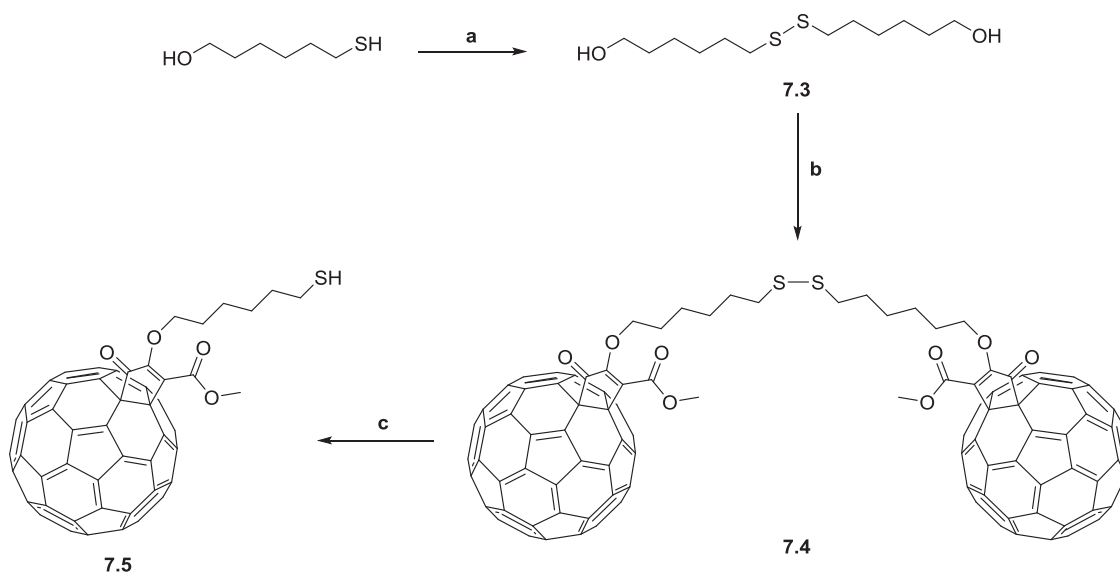
7.2.2 Second approach: zwitterion mediated functionalisation

After our first attempt, we explored a different approach devised by Zhang and co-workers.^[10] By reacting DMAP and DMAD with a double bond we could form a 5 membered ring on the fullerenes (Scheme 7.3).



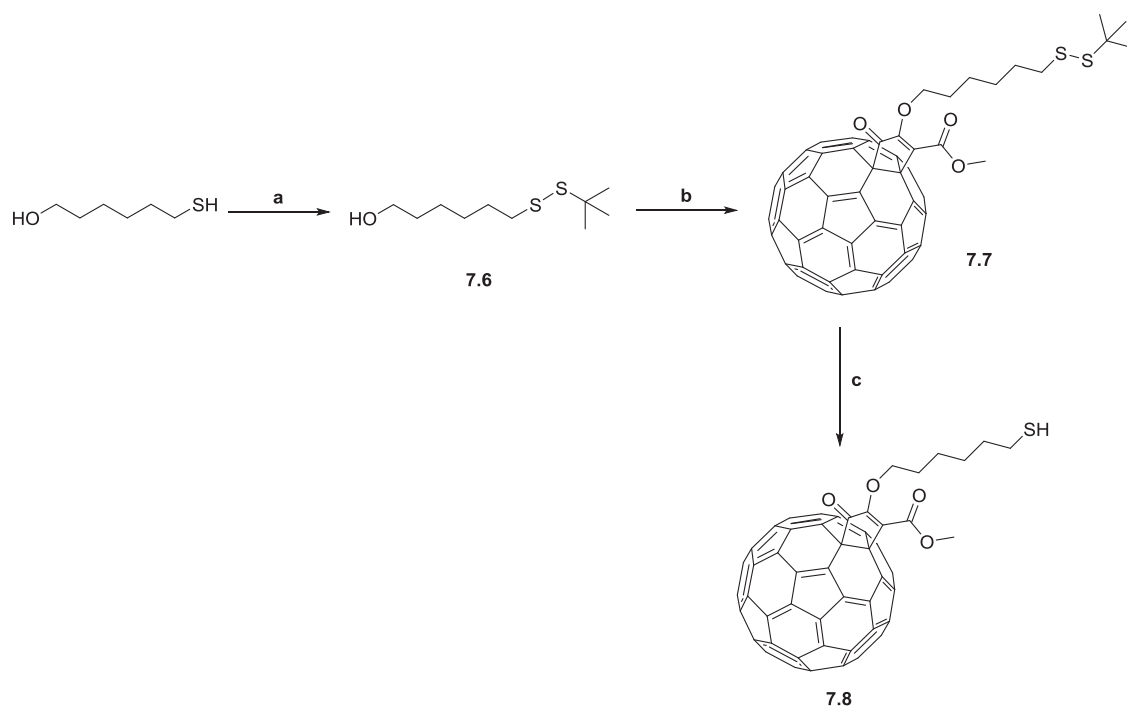
Scheme 7.3 Reaction mechanism of DMAD and DMAP on fullerenes.

By adding a different primary alcohol to the reaction we could trap the charged intermediate obtaining a substituted fullerene. A linear alkyne chain containing two functionalities, an alcohol on one end and thiol on the other, was firstly oxidised to the disulphide dimer before being coupled on C₆₀ (Scheme 7.4).



Scheme 7.4 Reagents and conditions: a) I₂, MeOH, r.t., 33%; b) DMAP, DMAD, C₆₀, toluene, r.t., 18%; c) DTT, DBU, DMF, r.t., quant..

The dimer **7.4** suffered from very poor solubility in any organic solvents; deprotection was quantitative on freshly made material, but after a short storage (both at r.t. or lower) the deprotection was harder to obtain as **7.4** tended to form indispersible aggregates. Looking at a more storable construct we tried to protect the thiol of the mercaptohexanol with a ^tBu group.^[11] Once coupled on the fullerenes the ^tBu could be removed with *tert*butylphosphine^[12] (Scheme 7.5).

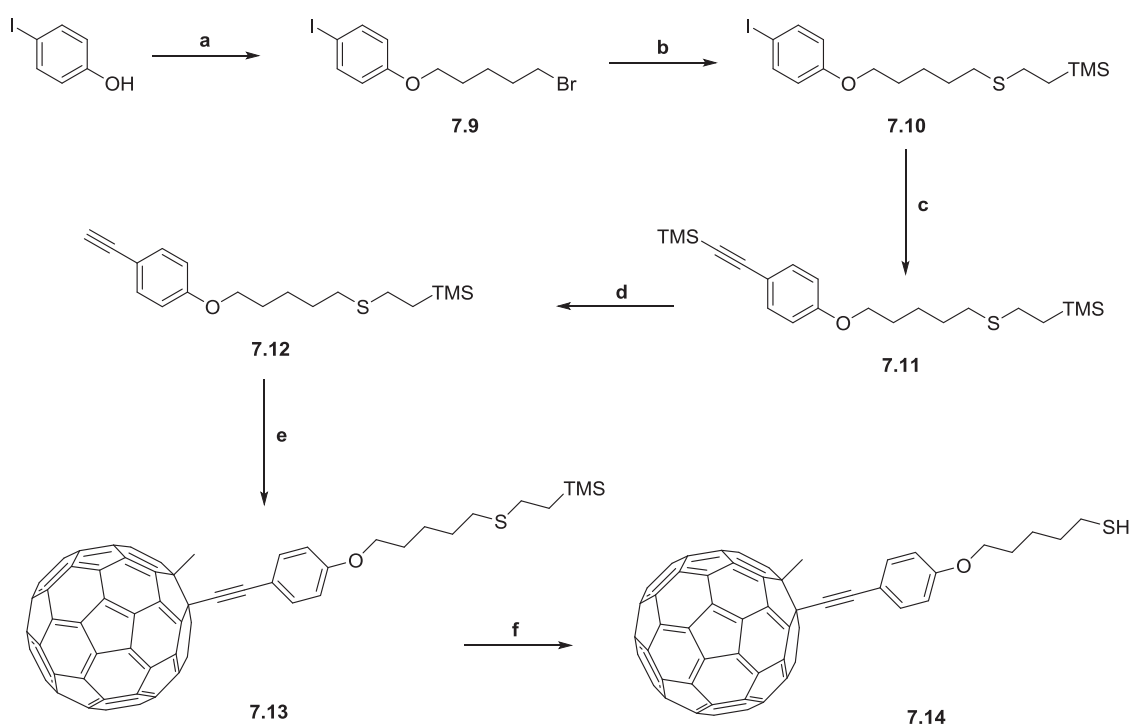


Scheme 7.5 Reagents and conditions: a) di-*tert*-butyl-1-(*tert*-butylthio)-1,2-hydrazinedicarboxylate, triethylamine, DMF, r.t., 46%; b) DMAP, DMAD, C₆₀, toluene, r.t., 67%; c) tributylphosphine, ODCB, r.t., 90%.

The protected precursor **7.7** presented better dispersibility, but solubility in organic solvents was still limited. Although the final thiofullerene **7.8** was successfully synthesised it presented some major drawbacks for its application in SPOS. First, it was difficult to keep the free thiol reduced while reacting with the first saccharide; second its solubility was very limited; and lastly, in some reaction conditions the 5 membered ring could possibly react through its conjugated double bond.

7.2.3 Third approach: functionalisation by ethynylation

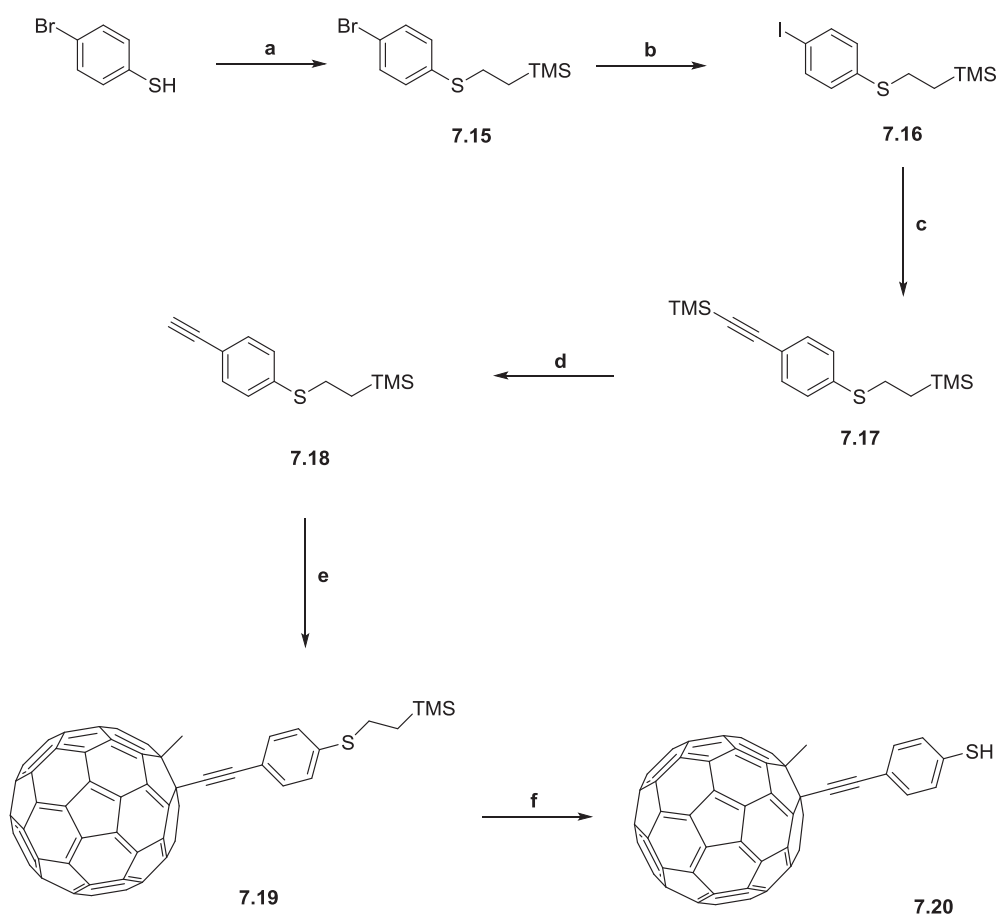
Finally, we opted for attaching the linker on the C₆₀ by nucleophilic attack. This strategy, described by Shirai et al. in 2006,^[13] is architecturally the most linear form of functionalisation we could perform on the fullerenes (Scheme 7.6).



Scheme 7.6 Reagents and conditions: a) 1,5-dibromopentane, TBAB, K₂CO₃, H₂O, 110 °C, 21%; b) TMS-ethanethiol, LHMDs, THF, r.t., 65%; c) CuI, triethylamine, TMS-acetylene, PdCl₂(PPh₃)₂, r.t., quant.; d) K₂CO₃, THF:MeOH (1:1), r.t., 34%; e) i. LHMDs, C₆₀, toluene, r.t.; ii. MeI, r.t., quant.; f) TBAF, toluene, r.t., 60%.

Once the alkyne attacks the double bond on the C₆₀ it leaves the other reactive end to be quenched. We tried both with TFA and MeI, the latter being a better substrate for the nucleophilic attack because the iodide is a prompt leaving group. Here again we obtained the desired thio-fullerene but its properties, in particular its solubility and stability as a free thiol, were still far from optimal for its future applications.

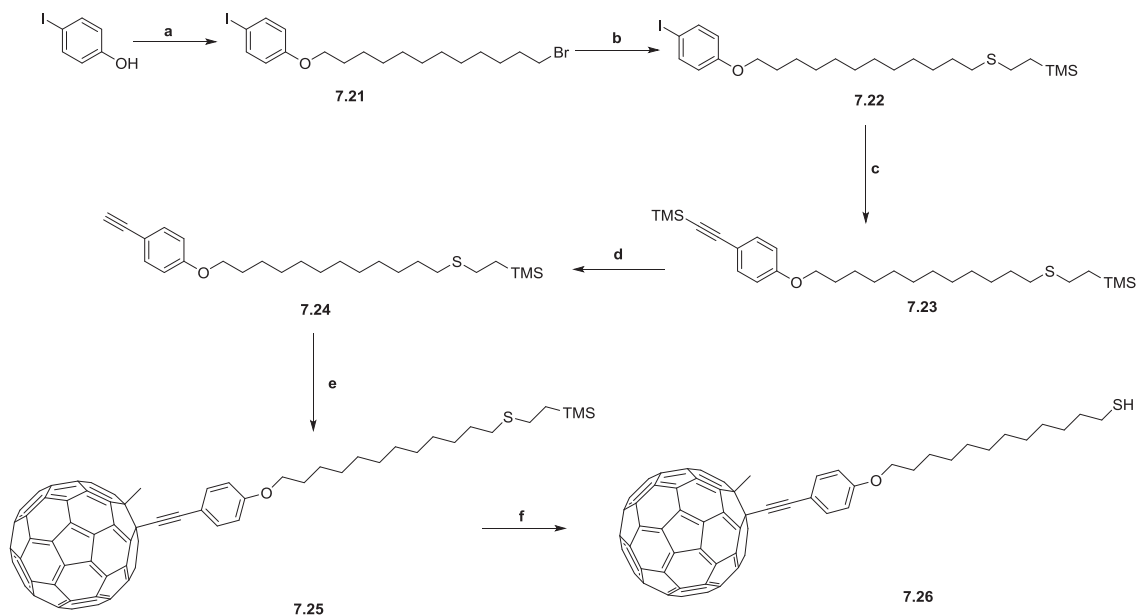
Considering that with a different type of functionalisation the same problems we encountered before persisted, we tried to modulate the physico-chemical properties of the thio-fullerene by changing the length of the alkyne chain, both decreasing and increasing the number of carbon atoms. In the first case we dispensed with the chain altogether (Scheme 7.7).



Scheme 7.7 Reagents and conditions: a) di-*tert*-butyl peroxide, vinyl-TMS, 100 °C, 34%; b) *t*BuLi, iodine, Et₂O, -78 °C, quant.; c) DIPEA, TMS-acetylene, CuI, PdCl₂(PPh₃)₂, 50 °C, 49%; d) K₂CO₃, THF:MeOH (1:1), r.t., 76%; e) i. LHMDS, C₆₀, toluene, r.t.; ii. MeI, r.t., quant.; f) TBAF, toluene, r.t., 49%.

We were hoping that the aromaticity of the substituent would have played a role in the solubility and stability of the fullerene-thiol. Unfortunately our approach did not yield the desired features.

Willing to explore the contribution of the lipophilicity of the alkyl chain, we extended it to 12 carbon atoms (Scheme 7.8).



Scheme 7.8 Reagents and conditions: a) 1,2-dibromododecane, NEt_4Br , K_2CO_3 , H_2O , 110°C , 37%; b) TMS-ethanethiol, LHMDS, THF, r.t., quant.; c) CuI , triethylamine, TMS-acetylene, $\text{PdCl}_2(\text{PPh}_3)_2$, r.t., quant.; d) K_2CO_3 , THF:MeOH (1:1), r.t., 78%; e) i. LHMDS, C_{60} , toluene, r.t.; ii. MeI , r.t., 18%; f) TBAF, toluene, r.t., 75%.

Once again, the solubility of the construct in organic solvents was not satisfactory and after a discussion with Prof. Shirai (see appendix) we decided to abandon the thio-derivatives as proposed tagging agent for SPOS.

7.3 Conclusions

In this brief chapter we intended to prepare and evaluate a thiol-derivative of C₆₀ fullerenes as support for SPOS. Fullerenes would allow solution-phase reactions with all the advantages of a solid-phase work-up, thanks to their solubility in a limited number of organic solvents. By installing a free thiol on the fullerenes we could form a disulphide with the first sugar of the saccharide chain we would like to synthesise, which would resist the subsequent glycosylation and deprotection steps and be promptly cleaved at the end of the reaction sequence.

After preparing thio-fullerenes using different types of functionalisation reactions and different lengths of the spacer, we agreed that these constructs would not be optimal as support for SPOS. Indeed, the thiols proved not only to be almost insoluble (poor solubility in ODCB or mixture of MeOH and toluene) but also too reactive to be handled in further reactions, forming un-dispersible di-fullerene disulphides.

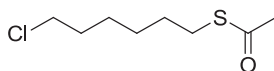
We concluded that future efforts should consider an alternative functional group to decorate the fullerenes onto which the growing saccharide chain will be appended.

7.4 References

- [1] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, 85, 2149-54.
- [2] S. J. Danishefsky and M. T. Bilodeau, *Angew. Chemie Int. Ed.* **1996**, 35, 1380-1419.
- [3] R. D. Guthrie, A. D. Jenkins, and J. Stehlicek, *J. Chem. Soc.* **1971**, 2690-96.
- [4] R. D. Guthrie, A. D. Jenkins, and G. A. F. Roberts, *J.C.S. Perkin I*, 2414-17.
- [5] O. J. Plante, *Science* **2001**, 291, 1523-27.
- [6] K. C. Nicolaou, N. Winssinger, J. Pastor, and F. DeRoose, *J. Am. Chem. Soc.* **1997**, 119, 449-50.
- [7] P. H. Seeberger and D. B. Werz, *Nature* **2007**, 446, 1046-51.
- [8] M. Maggini, G. Scorrano, and M. Prato, *J. Am. Chem. Soc.* **1993**, 115, 9798-99.
- [9] S. H. Kim, S. H. Lee, and S. H. Kang, *Tetrahedron Lett.* **1998**, 39, 9693-96.
- [10] W. Zhang, J. K. Sprafke, M. Ma, E. Y. Tsui, S. a Sydlik, G. C. Rutledge, and T. M. Swager, *J. Am. Chem. Soc.* **2009**, 131, 8446-54.
- [11] S. Far, C. Gouyette, and O. Melnyk, *Tetrahedron* **2005**, 61, 6138-42.
- [12] J. T. Ayers and S. R. Anderson, *Synth. Commun.* **1999**, 29, 351-58.
- [13] Y. Shirai, L. Cheng, B. Chen, and J. M. Tour, *J. Am. Chem. Soc.* **2006**, 128, 13479-89.

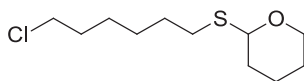
7.5 Supporting information

6-Chlorohexyl ethanethioate **7.1**^[1]



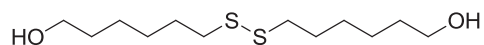
To a suspension of potassium thioacetate (1.24 g, 10.9 mmol, 1.1 eq.) in acetone (20 mL) was added 1-chloro-6-iodohexane (1.5 mL, 9.9 mmol, 1 eq.) and the mixture was stirred at r.t.. After 1 h more potassium thioacetate (120 mg, 1.05 mmol, 0.1 eq.) was added. After a further 3 h the reaction mixture was evaporated, recovered with water (150 mL) and extracted with DCM (3 x 150 mL). The organic layer was dried over MgSO₄ and the solvent was removed under vacuum affording the title compound **7.1** as a yellow oil (1.883 g, 9.67 mmol, 98%). **TLC** R_f = 0.75 (PE:EtOAc, 8:1); **¹H NMR** (CDCl₃, 400 MHz) δ 3.46 (t, J = 6.7 Hz, 2 H, CH₂Cl), 2.80 (t, J = 7.3 Hz, 2 H, CH₂S), 2.26 (s, 3 H, CH₃), 1.77-1.65 (m, 2 H, CH₂CH₂Cl), 1.52 (dt, J = 14.6, 7.4 Hz, 2 H, CH₂CH₂S), 1.44-1.26 (m, 4 H, ClCH₂CH₂CH₂CH₂CH₂CH₂S); **¹³C NMR** (CDCl₃, 101 MHz) δ 195.85 (CO), 44.92 (CH₂Cl), 32.38 (CH₂CH₂Cl), 30.62 (CH₃), 29.36 (CH₂CH₂S), 28.91 (CH₂S), 27.98 (CH₂CH₂CH₂S), 26.34 (CH₂CH₂CH₂Cl); **MS** (FI⁺) m/z = 195.73 [M+H]; **IR** (ATR) 2934.65 (CH), 1689.80 (CO).

6-Chlorohexyl thiotetrahydro-2H-pyran **7.2**^[2]



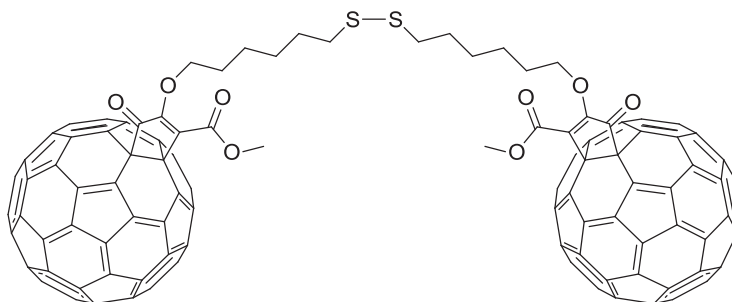
To a solution of **7.1** (937.6 mg, 4.81 mmol, 1 eq.) in MeOH (50 mL) was added $\text{BF}_3\text{Et}_2\text{O}$ (2.971 mL, 24 mmol, 5 eq.) and the mixture was stirred at r.t.. After 48 h DHP was added (2.197 mL, 24 mmol, 5 eq.). After 72 h the solvent was removed under vacuum. The residue was recovered with DCM (100 mL), washed with NaOH 0.1 M (100 mL) and the organic layer dried over MgSO_4 and evaporated. The crude was purified by flash column chromatography over alumina (DCM:PE, 0.5:9 to 1:1) to afford the title compound **7.2** as a white solid (456.3 mg, 1.93 mmol, 40%). **TLC** R_f = 0.8 (DCM:PE, 8:2); **^1H NMR** (CDCl_3 , 400 MHz) δ 3.65 (t, J = 6.0 Hz, 1 H, CH), 3.53 (t, J = 6.7 Hz, 2 H, CH_2Cl), 2.68, 2.55 (2 x dt, J = 12.4, 7.4 Hz, 2 x 1 H, OCH_2), 1.88-1.71 (m, 4 H, $\text{CH}_2\text{CH}_2\text{Cl}$, CH_2S), 1.65-1.54 (m, 6 H, $\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{S}$, SCHCH_2), 1.49-1.30 (m, 6 H, $\text{CH}_2\text{CH}_2\text{CH}_2\text{S}$, $\text{CHCH}_2\text{CH}_2\text{CH}_2$); **^{13}C NMR** (CDCl_3 , 63 MHz) δ 63.11 (CH), 52.42 (OCH_2), 45.48 (CH_2Cl), 36.24 ($\text{CH}_2\text{CH}_2\text{Cl}$), 32.88 (CH_2S), 32.61 ($\text{CH}_2\text{CH}_2\text{S}$), 30.32 ($\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{S}$), 29.65 ($\text{CH}_2\text{CH}_2\text{CH}_2\text{S}$), 28.71 (CHCH_2CH_2), 26.91 (OCH_2CH_2), 24.23 (CHCH_2); **MS** (FI^+) m/z = 237.81 [$\text{M}+\text{H}$]; **IR** (ATR) 2930.68 (CH).

Bis-(6-hydroxyhexyl) disulphide **7.3**^[3]



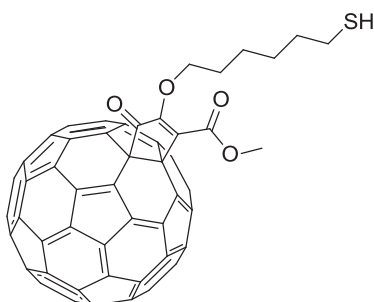
To a solution of 6-mercaptohexanol (500 μ L, 3.65 mmol, 1 eq.) in MeOH (1 mL) was added dropwise a 0.5 M solution of iodine in MeOH (3 mL), until the reaction turned to a persistent yellow. The reaction was stirred at r.t. for 0.5 h and then the excess of iodine quenched with a 10% aq. solution of sodium bisulphite, until the reaction became colourless. The reaction was recovered with H₂O (100 mL) and extracted with DCM (3 x 100 mL). The organic layer was dried over MgSO₄ anhydrous and the crude purified by flash column chromatography over silica (EtOAc:hexane, 1:1) to afford the title compound **7.3** as a colourless oil that crystallised on standing (321 mg, 1.2 mmol, 33%). **TLC** R_f = 0.35 (EtOAc:hexane, 1:1); **¹H NMR** (CDCl₃, 400 MHz) δ 3.55 (t, J = 6.6 Hz, 4 H, 2 x CH₂OH), 2.76 (s, 2 H, 2 x OH), 2.63 (t, J = 7.3 Hz, 4 H, 2 x CH₂S), 1.72-1.58 (m, 4 H, 2 x SCH₂CH₂), 1.58-1.44 (m, 4 H, 2 x CH₂CH₂OH), 1.44-1.25 (m, 8 H, 2 x SCH₂CH₂CH₂CH₂CH₂OH); **¹³C NMR** (CDCl₃, 101 MHz) δ 62.48 (2 x CH₂OH), 38.96 (2 x CH₂S), 32.45 (2 x CH₂CH₂OH), 29.08 (2 x SCH₂CH₂), 28.22 (2 x SCH₂CH₂CH₂), 25.35 (2 x CH₂CH₂CH₂OH); **MS** (ESI⁺) m/z = 267.47 [M+H]⁺; **IR** (ATR) 3429.50 (OH), 2925.93 (CH).

Bis-(6-thiohexyl fullerene) disulphide **7.4**



C₆₀ (10 mg, 0.014 mmol, 1 eq.) were sonicated with DMAP (54 mg, 0.44 mmol, 30 eq.) in toluene (8 mL). DMAD (3.4 μ L, 0.028 mmol, 2 eq.) was dissolved in toluene (0.5 mL) and added to the reaction dropwise. The reaction was stirred at r.t. for 30 min, then **7.3** (108 mg, 0.41 mmol, 30 eq.) was added. The reaction was stirred at r.t. for 16 h before the fullerenes were filtrated, washed with toluene and diethyl ether, to afford the title compound **7.4** as a brown solid (5 mg, 0.0026 mmol, 18%). **TLC** R_f = 0.1 (toluene, 100%); **¹H NMR** (toluene-d₈, 400 MHz) δ 3.77 (s, 6 H, 2 x CH₃), 3.48 (t, *J* = 6.3 Hz, 4 H, 2 x CH₂O), 2.44-2.39 (m, 4 H, 2 x CH₂S), 1.59-1.34 (m, 16 H, 2 x OCH₂CH₂CH₂CH₂CH₂CH₂S); **¹³C NMR** (toluene-d₈, 101 MHz) δ 202.51, 201.89 (2 x CO, 2 x COO), 158.89-141.22 (C-fullerene, 2 x COCCH₂), 114.43 (2 x CCOO), 69.87 (2 x CH₂O), 52.36 (2 x CH₃), 33.10 (2 x CH₂S), 31.92 (2 x CH₂CH₂O), 30.71 (2 x CH₂CH₂S), 25.52 (2 x CH₂CH₂CH₂O), 24.97 (2 x CH₂CH₂CH₂S); **MS** (LD⁺) *m/z* 1927.88 [M]; **IR** (ATR) 1667.71 (CO).

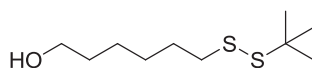
6-thiohexyl fullerene 7.5



7.4 (5 mg, 0.0026 mmol, 1 eq.) were dispersed in DMF (2 mL) by sonication. DTT (0.4 mg, 0.0026 mmol, 1 eq.) and DBU (0.4 μ L, 0.0026 mmol, 1 eq.) were added. The reaction was stirred at r.t. for 2 h then the solvent was removed under vacuum, the fullerenes sonicated in toluene and separated by centrifugation (3 x 10 mL), to afford the title compound **7.5** as a brown solid (5 mg, 0.0052 mmol, quant.); **TLC** R_f = 0.2 (hexane:EtOAc, 1:1); **MS** (LD^+) m/z 964.95 [M]; **IR** (ATR) 1674.13 (CO).

NMR spectra could not be obtained due to the poor solubility of the compound. After the deprotection the immediate formation of disulphides was observed.

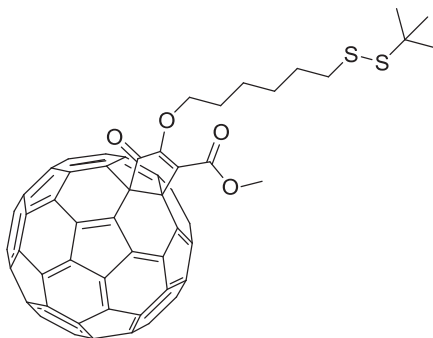
6-(*tert*-butyldisulphanyl) hexanol 7.6^[4]



To a solution of 6-mercaptohexanol (178 μ L, 1.3 mmol, 1 eq.) in DMF (5 mL) were added dropwise triethylamine (181.2 μ L, 1.3 mmol, 1 eq.) and di-*tert*-butyl-1-(*tert*-

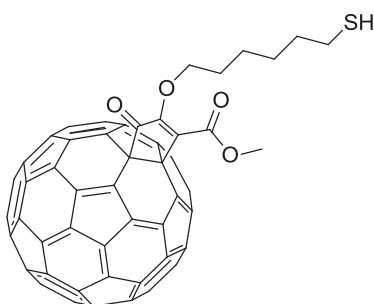
butylthio)-1,2-hydrazinedicarboxylate (500 mg, 1.56 mmol, 1.2 eq.). The reaction is stirred at r.t. for 24 h, before more di-*tert*-butyl-1-(*tert*-butylthio)-1,2-hydrazinedicarboxylate (250 mg, 0.78 mmol, 0.6 eq.) was added and the reaction was stirred for 16 h at r.t.. More triethylamine (362.4 μ L, 2.6 mmol, 2 eq.) was added and the reaction was stirred for another 24 h. The mixture was then evaporated and the crude purified by flash column chromatography over silica (EtOAc: hexane, 1:10 to 1:1) to afford the title compound **7.6** as a colorless oil that crystallised on standing (133.7 mg, 0.6 mmol, 46%). **TLC** R_f = 0.7 (EtOAc:hexane, 1:1); **^1H NMR** (CDCl_3 , 250 MHz) δ 3.66 (t, J = 6.5 Hz, 2 H, CH_2OH), 2.78-2.68 (m, 2 H, CH_2S), 1.99 (s, 1 H, OH), 1.79-1.55 (m, 4 H, $\text{CH}_2\text{CH}_2\text{OH}$, $\text{CH}_2\text{CH}_2\text{S}$), 1.47-1.38 (m, 4 H, $\text{SCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{OH}$), 1.35 (s, 9 H, ^tBu); **^{13}C NMR** (CDCl_3 , 63 MHz) δ 63.17 (CH_2OH), 48.12 ($\text{C}(\text{CH}_3)_3$), 41.19 (CH_2S), 32.96 ($\text{CH}_2\text{CH}_2\text{OH}$), 30.37 ($\text{C}(\text{CH}_3)_3$), 29.65 ($\text{CH}_2\text{CH}_2\text{S}$), 28.59 ($\text{CH}_2\text{CH}_2\text{CH}_2\text{OH}$), 25.79 ($\text{SCH}_2\text{CH}_2\text{CH}_2$); **MS** (FI^+) m/z = 222.11 [M]; **IR** (ATR) 3401.46 (OH), 2933.91 (CH).

6-(*tert*-butyldisulphanyl) hexyl fullerene **7.7**



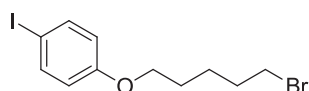
C₆₀ fullerenes (10 mg, 0.014 mmol, 1 eq.) were sonicated with DMAP (54 mg, 0.44 mmol, 30 eq.) in toluene (4.5 mL). DMAD (3.4 μ L, 0.028 mmol, 2 eq.) was dissolved in toluene (0.5 mL) and added to the reaction dropwise. The reaction was stirred at r.t. for 30 min, then **7.6** (123.7 mg, 0.56 mmol, 40 eq.) was dissolved in toluene (1 mL) and added to the reaction. The reaction was stirred at r.t. for 16 h before the solvent was removed under vacuum and the fullerenes dissolved in ODCB, precipitated with diethyl ether, filtered and washed with diethyl ether on a polyamido membrane (0.2 μ m pore size), to afford the title compound **7.7** as a brown solid (10 mg, 0.0095 mmol, 67%); **TLC** R_f = 0.2 (toluene, 100%); **¹H NMR** (toluene-d₈, 400 MHz) δ 3.81 (s, 3 H, CH₃), 3.51 (t, *J* = 6.3 Hz, 2 H, CH₂O), 2.48-2.41 (m, 2 H, CH₂S), 1.59-1.34 (m, 8 H, OCH₂CH₂CH₂CH₂CH₂CH₂S), 0.04 (s, 9 H, ^tBu); **¹³C NMR** (toluene-d₈, 101 MHz) δ 202.98, 202.76 (CO, COO), 158.62-140.99 (C-fullerene, COCCH₂), 112.75 (CCOO), 70.55 (CH₂O), 50.11 (CH₃), 35.87 (CH₂S), 33.49 (CH₂CH₂O), 31.77 (CH₂CH₂S), 26.98 (CH₂CH₂CH₂O), 23.22 (CH₂CH₂CH₂S); **MS** (LD⁺) *m/z* 1053.12 [M]; **IR** (ATR) 1682.99 (CO).

6-thiohexyl fullerene **7.8**



7.7 (10 mg, 0.0095 mmol, 1 eq.) was dispersed in ODCB (4 mL) by sonication. Tributylphosphine (12 μ L, 0.048 mmol, 5 eq.) and a few drops of H₂O were added. The reaction was stirred at r.t. for 18 h, recovered with DMF, precipitated with Et₂O, filtered on a polyamido membrane (0.2 μ m pore size) and washed with more Et₂O to afford the title compound **7.8** as a brown solid (8 mg, 90%); **TLC** R_f = 0.2 (hexane:EtOAc, 1:1); **MS** (LD⁺) m/z 964.95 [M]; **IR** (ATR) 1687.89 (CO). NMR spectra could not be obtained due to the poor solubility of the compound. After the deprotection the immediate formation of disulphides was observed.

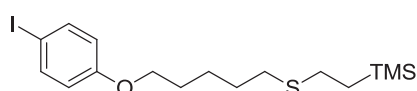
(5-bromopentyl)oxy-4-iodobenzene 7.9^[5]



To a solution of 4-iodophenol (1.12 g, 5.1 mmol, 1 eq.) in H₂O (10 mL), were added K₂CO₃ (704 mg, 5.1 mmol, 1 eq.), 1,5-dibromopentane (832 μ L, 6.1 mmol, 1.2 eq.) and tetrabutylammonium bromide (164 mg, 0.51 mmol, 1 eq.). The reaction mixture was stirred for 48 h refluxing at 110 °C, before being recovered with H₂O (150 mL) and extracted with DCM (3 x 150 mL). The organic layer was dried over MgSO₄ and evaporated, and the crude purified by flash column chromatography over silica (DCM:PE, 1:10) to afford the title compound **7.9** as a white solid (397.5 mg, 1.08 mmol, 21%). **TLC** R_f = 0.3 (DCM:PE, 1:10); **¹H NMR** (CDCl₃, 400 MHz) δ 7.54 (d, *J* = 9.0 Hz, 2 H, 2 x CHCl-arom), 6.66 (d, *J* = 9.0 Hz, 2 H, 2 x CHCO-arom), 3.91 (t, *J* = 6.3 Hz, 2 H, CH₂O), 3.43 (t, *J* = 6.8 Hz, 2 H, CH₂Br), 2.04-1.84 (m, 2 H, CH₂CH₂O), 1.84-1.71 (m, 2 H, CH₂CH₂Br), 1.68-1.50

(m, 2 H, CH₂CH₂CH₂O); ¹³C NMR (CDCl₃, 101 MHz) δ 158.88 (CO-arom), 138.51 (2 x CHCl-arom), 116.98 (2 x CHCO-arom), 82.72 (Cl-arom), 67.73 (CH₂O), 33.70 (CH₂Br), 32.51 (CH₂CH₂O), 28.39 (CH₂CH₂Br), 24.85 (CH₂CH₂CH₂CO); MS (FI⁺) m/z = 367.93 [M]; IR (ATR) 3096.97 (C=C arom), 2941.17 (CH).

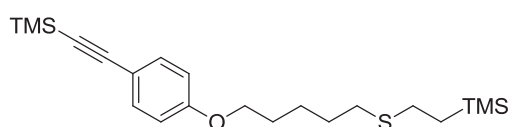
5-(4-iodophenoxy)-pentyl thioethyltrimethylsilane **7.10**



To a solution of trimethyl-silyl-ethanethiol (175 μL, 1.09 mmol, 1 eq.) in THF (dry, 2.5 mL) was added, dropwise at 0 °C, a solution of LHMDs (186 mg, 1.11 mmol, 1.02 eq.) in THF (dry, 2 mL). The reaction was stirred for 5 min, then a solution of **7.9** (397.5 mg, 1.09 mmol, 1 eq.) in THF (dry, 2.5 mL) was added dropwise. The reaction mixture was stirred for 16 h at r.t., then quenched with H₂O and diluted with hexane (10 mL). The aqueous phase was extracted with DCM (3 x 50 mL). The organic layers were combined, dried over MgSO₄ and evaporated. The crude was purified by flash column chromatography over silica (DCM:PE, 1:10) to afford the title compound **7.10** as a white solid (295.3 mg, 0.7 mmol, 65%). TLC R_f = 0.2 (DCM:PE, 1:10); ¹H NMR (CDCl₃, 400 MHz) δ 7.53 (d, *J* = 9.0 Hz, 2 H, 2 x CHCl-arom), 6.66 (d, *J* = 9.0 Hz, 2 H, 2 x CHCO-arom), 3.91 (t, *J* = 6.4 Hz, 2 H, CH₂O), 2.55 (t, *J* = 8.7 Hz, 4 H, SCH₂CH₂Si), 1.89-1.70 (m, 2 H, CH₂CH₂O), 1.70-1.59 (m, 2 H, CH₂CH₂S), 1.59-1.47 (m, 2 H, CH₂CH₂CH₂O), 0.99-0.77 (m, 2 H, CH₂S), 0.02 (s, 9 H, ^tBu); ¹³C NMR (CDCl₃, 63 MHz) δ 168.33 (CO-arom), 137.12 (2 x CHCC-arom), 114.66 (2 x CHCO-arom), 112.19 (Cl-

arom), 67.35 (CH₂O), 37.42 (CH₂Si), 33.01 (CH₂CH₂S), 30.00 (CH₂CH₂O), 27.44 (CH₂CH₂Si), 25.22 (CH₂CH₂CH₂O), 18.99 (CH₂S), 0.07 (^tBu); **MS** (FI⁺) *m/z* = 422.44 [M]; **HRMS** (FI⁺) found 422.0600 [M], C₁₆H₂₇IOSSi requires 422.0597; **IR** (ATR) 3044.53 (C=C arom), 2943.95 (CH).

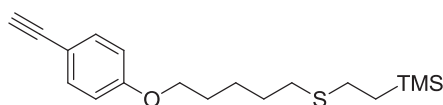
5-(4-trimethylsilyl ethynylphenoxy)-pentyl thioethyltrimethylsilane **7.11**



To a solution of **7.10** (295 mg, 0.70 mmol, 1 eq) in triethylamine (5 mL), were added CuI (4 mg, 0.021 mmol, 0.03 eq.), trimethylsilyl acetylene (149 μ L, 1.05 mmol, 1.5 eq.) and PdCl₂(PPh₃)₂ (10 mg, 0.014 mmol, 0.02 eq.). The reaction was stirred for 48 h at r.t., then quenched with a satd. solution of NH₄Cl. The aqueous phase was extracted with DCM (3 x 150 mL). The organic layer was dried over MgSO₄ and evaporated to afford the title compound **7.11** as brown oil (282 mg, 0.7 mmol, quant.). **TLC** *R_f* = 0.15 (DCM:PE, 1:10); **¹H NMR** (CDCl₃, 400 MHz) δ 7.38 (d, *J* = 8.9 Hz, 2 H, 2 x CHCC-arom), 6.79 (d, *J* = 8.9 Hz, 2 H, 2 x CHCO-arom), 3.94 (t, *J* = 6.5 Hz, 2 H, CH₂O), 2.67-2.42 (m, 4 H, SCH₂CH₂Si), 1.90-1.72 (m, 2 H, CH₂CH₂O), 1.71-1.60 (m, 2 H, CH₂CH₂S), 1.60-1.50 (m, 2 H, CH₂CH₂CH₂O), 0.96-0.74 (m, 2 H, CH₂S), 0.23 (s, 9 H, ^tBu-alkyne), 0.02 (s, 9 H, ^tBu-TMS); **¹³C NMR** (CDCl₃, 63 MHz) δ 167.91 (CO-arom), 136.04 (2 x CHCC-arom), 117.91 (2 x CHCO-arom), 110.21 (CC-arom), 98.97, 98.42 (CC-alkyne), 68.22 (CH₂O), 36.11 (CH₂Si), 33.94 (CH₂CH₂S), 31.68 (CH₂CH₂O), 27.10

(CH₂CH₂Si), 24.19 (CH₂CH₂CH₂O), 17.67 (CH₂S), 0.07, 0.04 (2 x ^tBu); **MS** (FI⁺) m/z = 392.75 [M]; **HRMS** (FI⁺) found 392.2026 [M], C₂₁H₃₆OSSi₂ requires 392.2025; **IR** (ATR) 3072.43 (C=C arom), 2941.18 (CH), 2212.42 (alkyne).

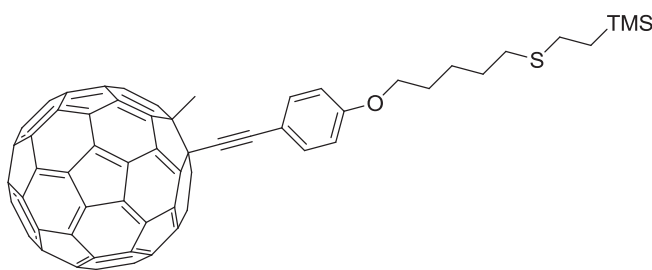
5-(4-ethynylphenoxy)-pentyl thioethyltrimethylsilane **7.12**



To a solution of **7.11** (282 mg, 0.72 mmol, 1 eq.) in MeOH:THF (1:1, 30 mL), was added K₂CO₃ (150 mg, 1.08 mmol, 1.5 eq.). The reaction was stirred for 1.5 h at r.t., then diluted with H₂O and hexane. The aqueous phase was extracted with DCM (3 x 150 mL). The organic layer was dried over MgSO₄ and evaporated, the crude filtered over silica (DCM:PE, 1:1) to afford the title compound **7.12** as a yellow oil (78 mg, 0.24 mmol, 34%). **TLC** R_f = 0.2 (DCM:PE, 1:10); **¹H NMR** (CDCl₃, 400 MHz) δ 7.41 (d, *J* = 8.8 Hz, 2 H, 2 x CHCC-arom), 6.82 (d, *J* = 8.9 Hz, 2 H, 2 x CHCO-arom), 3.95 (t, *J* = 6.4 Hz, 2 H, CH₂O), 2.99 (s, 1 H, CH-alkyne), 2.55 (t, *J* = 8.7 Hz, 4 H, SCH₂CH₂Si), 1.91-1.72 (m, 2 H, CH₂CH₂O), 1.72-1.61 (m, 2 H, CH₂CH₂S), 1.61-1.46 (m, 2 H, CH₂CH₂CH₂O), 1.01-0.74 (m, 1 H, CH₂S), 0.02 (s, 9 H, ^tBu); **¹³C NMR** (CDCl₃, 63 MHz) δ 161.18 (CO-arom), 135.31 (2 x CHCC-arom), 116.17 (2 x CHCO-arom), 115.71 (CC-arom), 85.47 (CH-alkyne), 77.45 (CCH-alkyne), 69.54 (CH₂O), 33.51 (CH₂Si), 31.05 (CH₂CH₂S), 30.55 (CH₂CH₂O), 29.42 (CH₂CH₂Si), 27.17 (CH₂CH₂CH₂O), 19.13 (CH₂S), -0.01 (^tBu); **MS** (FI⁺) m/z = 320.56 [M]; **HRMS** (FI⁺) found 320.1636 [M],

C₁₈H₂₀OSSi requires 320.1630; **IR** (ATR) 3290.58 (CCH-alkyne), 3090.44 (C=C arom), 2945.73 (CH) , 2208.39 (alkyne).

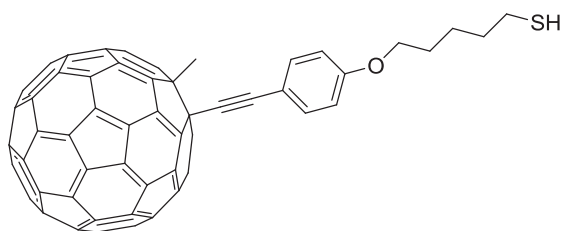
Ethynylphenoxy-4-(pentyl thioethyltrimethylsilane)-methyl fullerene **7.13**



To a solution of **7.12** (714.7 mg, 2.23 mmol, 1 eq.) in toluene (dry, 20 mL) was added LHMDS (1 M in THF, 4.46 mL, 4.46 mmol, 2 eq.), stirred for a few min and then added a solution of C₆₀ (1.9 g, 2.67 mmol, 1.2 eq.) in toluene (dry, 240 mL). The reaction was stirred at r.t. for a few min then quenched with MeI (2.776 mL, 44.6 mmol, 20 eq). The reaction mixture was evaporated and purified by precipitation with PE from toluene to afford the title compound **7.13** as brown solid (2.593 g, 2.2 mmol, quant.). **TLC** R_f = 0.25 (toluene:MeOH, 8:2); ¹H NMR (toluene-d₈, 400 MHz) δ 7.74 (d, *J* = 8.6 Hz, 2 H, 2 x CHCC-arom), 6.82 (d, *J* = 8.7 Hz, 2 H, 2 x CHCO-arom), 6.77 (s, 3 H, CH₃), 3.57 (t, *J* = 6.3 Hz, 2 H, CH₂O), 2.60-2.47 (m, 2 H, CH₂CH₂Si), 2.45-2.37 (m, 2 H, CH₂S), 1.63-1.27 (m, 6 H, OCH₂CH₂CH₂CH₂CH₂S), 0.97-0.78 (m, 2 H, CH₂Si), -0.02 (s, 9 H, ^tBu); ¹³C NMR (toluene-d₈, 101 MHz) δ 162.37 (CO-arom), 154.16-144.44 (C-fullerene), 144.36 (2 x CHCC-arom), 127.68 (CC-arom), 117.19 (2 x CHCO-arom), 93.64, 86.32 (CC-alkyne), 69.99 (CH₂O), 64.21 (CH₃), 34.19 (CH₂S), 31.76 (CH₂CH₂O), 31.29

(CH₂CH₂S), 30.05 (CH₂CH₂Si), 27.81 (CH₂CH₂CH₂O), 19.63 (CH₂Si), 0.35 (^tBu); **MS** (LD⁺) m/z 1056.24 [M]; **IR** (ATR) 3093.11 (C=C arom), 2944.77 (CH) , 2201.57 (alkyne).

Ethynylphenoxy-4-(pentyl thiol)-methyl fullerene **7.14**

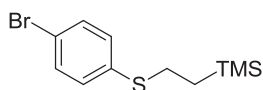


To a solution of **7.13** (30 mg, 0.028 mmol, 1 eq.) in toluene (dry, 30 mL) was added TBAF (1 M in THF, 1 mL, 1 mmol, 36 eq.) and stirred at r.t. for 2 h. The reaction was quenched with MeOH (30 mL), and TFA (200 μ L). The reaction mixture was evaporated and purified by precipitation with MeOH to afford the title compound **7.14** as a black solid (16.1 mg, 0.016 mmol, 60%). **TLC** R_f = 0.0 (toluene:MeOH, 8:2); **MS** (LD⁺) m/z 956.00 [M]; **IR** (ATR) 3091.37 (C=C arom), 2940.55 (CH) , 2200.92 (alkyne).

NMR spectra could not be obtained due to the poor solubility of the compound.

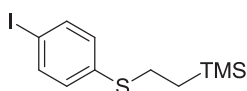
After the deprotection the immediate formation of disulphides was observed.

2-trimethylsilylethyl-4-bromophenyl sulphide **7.15**^[6]



To 4-bromothiophenol (6.45 g, 34.1 mmol, 1 eq.) were added di-*tert*-butyl peroxide (900 μ L, 4.78 mmol, 0.14 eq.) and vinyl trimethylsilane (6 mL, 40.9 mmol, 1.2 eq.). The reaction was stirred at 100 °C for 16 h, then recovered with hexane, washed with 10% NaOH (200 mL), dried over MgSO₄ and evaporated. The crude was purified by flash column chromatography over silica (hexane, 100%) to afford the title compound **7.15** as colourless oil (3.3448 g, 11.6 mmol, 34%). **TLC** R_f = 0.15 (hexane, 100%); **¹H NMR** (CDCl₃, 400 MHz) δ 7.33 (d, *J* = 8.5 Hz, 2 H, 2 x CHCBr), 7.11 (d, *J* = 8.5 Hz, 2 H, 2 x CHCS), 2.96-2.82 (m, 2 H, SCH₂), 0.97-0.78 (m, 2 H, CH₂TMS), -0.00 (s, 9 H, TMS); **¹³C NMR** (CDCl₃, 101 MHz) δ 160.11 (CS), 137.97 (2 x CHCBr), 116.52 (2 x CHCO), 115.09 (CBr), 27.87 (SCH₂), 19.29 (CH₂TMS), 0.00 (TMS); **MS** (FI⁺) *m/z* = 289.31 [M]; **IR** (ATR) 3092.17 (C=C arom), 2946.15 (CH).

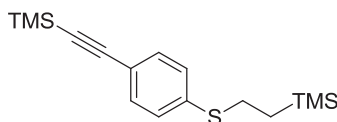
2-trimethylsilylethyl-4-iodophenyl sulphide **7.16**^[6]



To a solution of **7.15** (3.3448 g, 11.56 mmol, 1 eq.) in Et₂O (dry, 75 mL) was added *t*-BuLi (1.7 M in pentane, 15 mL, 25.5 mmol, 2.2 eq.) at -78 °C and stirred

for 1 h. Iodine (3.8 g, 15 mmol, 1.3 eq.) was dissolved in Et₂O (dry, 75 mL), cooled to -78 °C and added to the reaction. The reaction was stirred at -78 °C for 10 min, allowed to warm up to 0 °C and stirred for another 30 min. The reaction was washed with satd. solution of Na₂S₂O₃ until the disappearance of the iodine. The organic layer was dried over MgSO₄ and evaporated to afford the title compound **7.16** as a yellow oil (3.7553 g, 11.16 mmol, 97%). **TLC** R_f = 0.2 (hexane, quant.); **¹H NMR** (CDCl₃, 400 MHz) δ 7.60-7.41 (m, 1 H, 2 x CHCl), 7.04-6.88 (m, 2 H, 2 x CHCS), 3.00-2.70 (m, 2 H, SCH₂), 0.96-0.76 (m, 2 H, CH₂TMS), -0.00 (s, 9 H, TMS); **¹³C NMR** (CDCl₃, 101 MHz) δ 159.08 (CS), 136.44 (2 x CHCl), 116.48 (2 x CHCO), 115.13 (Cl), 27.99 (SCH₂), 19.36 (CH₂TMS), -0.00 (TMS); **MS** (FI⁺) m/z = 336.31 [M]; **IR** (ATR) 3091.99 (C=C arom), 2946.55 (CH).

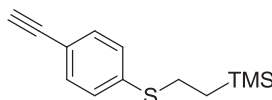
2-trimethylsilylethyl-4-trimethylsilylethynylphenyl sulphide **7.17**^[6]



To a solution of **7.16** (3.7553 g, 11.16 mmol, 1 eq) in DIPEA (200 mL) were added trimethylsilyl acetylene (1.8 mL, 12.72 mmol, 1.1 eq), Pd(PPh₃)₂Cl₂ (244 mg, 0.35 mmol, 0.03 eq) and CuI (242 mg, 1.27 mmol, 0.11 eq). The reaction was stirred at 50 °C for 2 h, then evaporated and the crude purified by flash column chromatography over silica (DCM:hexane, 0:10 to 1:10) to afford the title compound **7.17** as a yellow oil (1.6813 g, 5.48 mmol, 49%). **TLC** R_f = 0.4 (DCM:hexane, 1:10); **¹H NMR** (CDCl₃, 400 MHz) δ 7.37 (d, *J* = 8.5 Hz, 2 H, 2 x

CHCC), 7.18 (d, J = 8.5 Hz, 2 H, 2 x CHCS), 3.08-2.81 (m, 2 H, SCH₂), 1.12-0.73 (m, 2 H, CH₂TMS), 0.24 (s, 9 H, alkyne-TMS), 0.04 (s, 9 H, TMS); ¹³C NMR (CDCl₃, 101 MHz) δ 160.34 (CS), 136.72 (2 x CHCC), 116.69 (2 x CHCO), 116.02 (CCCTMS), 85.81 (CCTMS), 77.73 (CTMS), 27.89 (SCH₂), 19.22 (CH₂TMS), 0.03 (C-alkyne-TMS), -0.00 (CH₂TMS); MS (EI⁺) m/z = 306.61 [M]; IR (ATR) 3090.52 (C=C arom), 2946.15 (CH), 2203.76 (alkyne).

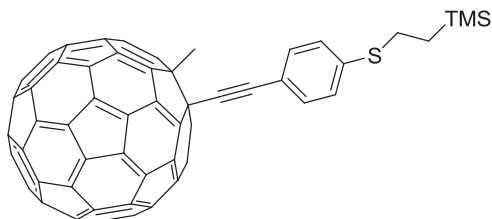
2-trimethylsilylethyl-4-ethynylphenyl sulphide **7.18**^[6]



To a solution of **7.17** (1.6813 g, 5.48 mmol, 1 eq.) in MeOH:THF (1:1, 25 mL), was added K₂CO₃ (122 mg, 0.88 mmol, 0.16 eq.). The reaction was stirred at r.t. for 9 h, recovered with DCM (150 mL) and washed with water (150 mL). The organic layer was dried over MgSO₄, evaporated and the residue purified by flash column chromatography over silica (DCM:hexane, 0:10 to 1:10) to afford the title compound **7.18** as a yellow oil (981.2 mg, 4.18 mmol, 76%). TLC R_f = 0.35 (DCM:hexane, 1:10); ¹H NMR (CDCl₃, 400 MHz) δ 7.39 (d, J = 8.4 Hz, 2 H, 2 x CHCC), 7.21 (d, J = 8.4 Hz, 2 H, 2 x CHCS), 3.08 (s, 1 H, CH-alkyne), 3.02-2.87 (m, 2 H, SCH₂), 1.02-0.86 (m, 2 H, CH₂TMS), 0.05 (s, 9 H, TMS); ¹³C NMR (CDCl₃, 101 MHz) δ 160.44 (CS), 135.89 (2 x CHCC), 116.57 (2 x CHCO), 116.38 (CCCH), 86.04 (CCH-alkyne), 77.92 (CH-alkyne), 27.56 (SCH₂), 19.39

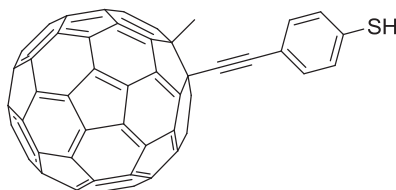
(CH₂TMS), -0.00 (CH₂**TMS**); **MS** (FI⁺) m/z = 234.43 [M]; **IR** (ATR) 3291.17 (CCH-alkyne), 3089.14 (C=C arom), 2944.77 (CH) , 2207.19 (alkyne).

Ethynylphenoxy-4-(thioethyltrimethylsilane)-methyl fullerene **7.19**



To a solution of **7.18** (250 mg, 0.75 mmol, 1 eq.) in toluene (20 mL) was added LHMDS (1 M in THF, 1.8 mL, 1.8 mmol, 1.2 eq.). The reaction was stirred at r.t. for 15 min, then a solution of C₆₀ (1.1 g, 1.5 mmol, 2 eq.) in toluene (80 mL) was added. After 15 min the reaction was quenched with MeI (940 μ L, 15 mmol, 20 eq.), and stirred at r.t. for other 3 h. The reaction was evaporated, the product precipitated from toluene with PE, hexane and filtered over silica with toluene to afford the title compound **7.19** as black solid (309.5 mg, 0.32 mmol, 42%). **TLC** R_f = 0.9 (toluene:DCM:hexane, 2:2:6); **¹H NMR** (toluene-d₈, 400 MHz) δ 7.60-7.20 (m, 4 H, 4 x CH-arom), 3.10-2.36 (m, 2 H, SCH₂), 1.34 (s, 3 H, CH₃), 1.02-0.68 (m, 2 H, CH₂TMS), -0.08 (s, 9 H, TMS); **¹³C NMR** (CDCl₃, 126 MHz) δ 156.41-119.98 (C-fullerene, C-arom, CH-arom), 79.65 (C-alkyne-C₆₀), 77.44 (C-alkyne-Ar), 29.50 (CH₃), 29.13 (SCH₂), 17.12 (CH₂TMS), -1.12 (TMS); **MS** (LD⁺) m/z 969.10 [M]; **IR** (ATR) 3097.16 (C=C arom), 2934.57 (CH) , 2200.01 (alkyne).

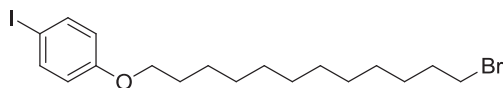
Ethynylphenoxy-4-(thio)-methyl fullerene **7.20**



To a solution of **7.19** (30.95 mg, 0.032 mmol, 1 eq.) in toluene (dry, 30 mL) was added TBAF (1 M in THF, 1 mL, 1 mmol, 30 eq.) and stirred at r.t. for 2 h. The reaction was quenched with MeOH (30 mL), and TFA (200 μ L). The reaction mixture was evaporated and purified by precipitation with MeOH to afford the title compound **7.20** as a black solid (19 mg, 0.022 mmol, 69%). **TLC** R_f = 0.0 (toluene:MeOH, 8:2); **MS** (LD^+) m/z 868.87 [M]; **IR** (ATR) 3091.24 (C=C arom), 2942.36 (CH) , 2204.47 (alkyne).

NMR spectra could not be obtained due to the poor solubility of the compound. After the deprotection the immediate formation of disulphides was observed.

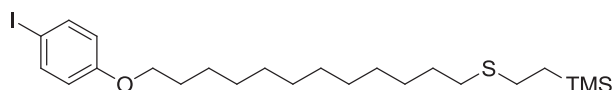
(12-bromododecyl)oxy-4-iodobenzene **7.21**^[7]



To a solution of 4-iodophenol (4.48 g, 20.4 mmol, 1 eq.) in H₂O (40 mL), were added 1,12-dibromododecane (8.02 g, 24.44 mmol, 1.2 eq.), K₂CO₃ (2.8 g, 20.4 mmol, 1 eq.) and NEt₄Br (645 mg, 2 mmol, 0.1 eq.). The reaction was stirred at

110 °C for 2 days, recovered with H₂O (100 mL) and extracted with DCM (3 x 150 mL). The organic layer was dried over MgSO₄, evaporated, and the residue purified by flash column chromatography over silica (DCM:hexane, 0:10 to 10:0) to afford the title compound **7.21** as white solid (3.5996 g, 7.7 mmol, 37%). **TLC** R_f = 0.4 (hexane:DCM, 10:1); **¹H NMR** (CDCl₃, 400 MHz) δ 7.54 (d, *J* = 8.9 Hz, 2 H, 2 x CHCl), 6.67 (d, *J* = 8.9 Hz, 2 H, 2 x CHCO), 3.91 (t, *J* = 6.6 Hz, 2 H, OCH₂), 3.41 (t, *J* = 6.9 Hz, 2 H, CH₂Br), 1.90-1.81 (m, 2 H, CH₂CH₂Br), 1.81-1.70 (m, 2 H, OCH₂CH₂), 1.46-1.28 (m, 16 H, 8 x CH₂); **¹³C NMR** (CDCl₃, 101 MHz) δ 160.47 (CO), 135.72 (2 x CHCl), 115.02 (2 x CHCO), 69.99 (OCH₂), 33.36 (CH₂CH₂CH₂Br), 31.39 (CH₂CH₂CH₂Br), 32.07, 31.83, 31.77, 30.75, 30.28, 30.11, 30.04, 29.54 (8 x CH₂), 29.48 (OCH₂CH₂); **MS** (EI⁺) *m/z* = 467.22 [M]; **IR** (ATR) 3091.47 (C=C arom), 2944.99 (CH).

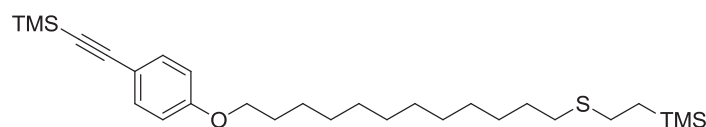
12-(4-iodophenoxy)-dodecyl thioethyltrimethylsilane **7.22**



To a solution of trimethylsilylethanethiol (1.3 mL, 8.16 mmol, 1.03 eq.) in THF (15 mL), was added LHMDs (1 M in THF, 8 mL, 1.01 eq.) at 0 °C, dropwise. After 5 min **7.21** (3.7 g, 7.92 mmol, 1 eq.) was added and the reaction was stirred at r.t. for 16 h. The reaction was quenched with water, diluted with hexane and the aqueous layer extracted with DCM (125 mL x 3). The organic layer was dried over MgSO₄ and evaporated to afford the title compound **7.22** as a yellow oil (4.95 g, 8.1 mmol, quant.). **TLC** R_f = 0.3 (hexane:DCM, 10:1); **¹H NMR** (CDCl₃,

400 MHz) δ 7.53 (d, J = 8.9 Hz, 2 H, 2 x CHCl), 6.66 (d, J = 8.9 Hz, 2 H, 2 x CHCO), 3.90 (t, J = 6.6 Hz, 2 H, OCH₂), 2.59-2.47 (m, 4 H, CH₂SCH₂), 1.80-1.70 (m, 2 H, OCH₂CH₂), 1.64-1.52 (m, 2 H, CH₂CH₂CH₂S), 1.47-1.21 (m, 16 H, 8 x CH₂), 0.90 -0.79 (m, 2 H, CH₂TMS), 0.02 (s, 9 H, TMS); ¹³C NMR (CDCl₃, 101 MHz) δ 160.89 (CO), 135.82 (2 x CHCl), 115.26 (2 x CHCO), 69.34 (OCH₂), 33.67 (CH₂CH₂CH₂S), 31.83 (CH₂CH₂CH₂S), 31.52, 31.42, 31.84, 30.95, 30.77, 30.49, 30.22, 29.87 (8 x CH₂), 29.48 (OCH₂CH₂), 27.62 (SCH₂CH₂TMS), 19.04 (CH₂TMS), -0.00 (CH₂TMS); MS (EI⁺) m/z = 520.63 [M]; IR (ATR) 3090.87 (C=C arom), 2943.52 (CH).

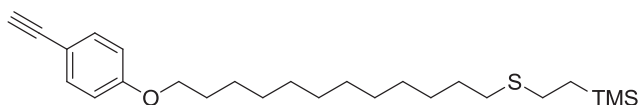
12-(4-ethynyltrimethylsilyl phenoxy)-dodecyl thioethyltrimethylsilane 7.23



To a solution of **7.22** (4.95 g, 7.92 mmol, 1 eq) in triethylamine (50 mL), were added TMSA (1.7 mL, 11.88 mmol, 1.5 eq), Pd(PPh₃)₂Cl₂ (95 mg, 0.14 mmol, 0.017 eq) and CuI (45 mg, 0.24 mmol, 0.03 eq). The reaction was stirred at r.t. for 16 h, then quenched with NH₄Cl solution, diluted with hexane and washed with more NH₄Cl solution. The organic layer was dried over MgSO₄ and evaporated to afford the title compound **7.23** as a yellow oil (4.99 g, 7.9 mmol, quant). TLC R_f = 0.35 (hexane:DCM, 10:1); ¹H NMR (CDCl₃, 400 MHz) δ 7.40-7.34 (m, 2 H, 2 x CHCC), 6.80 (d, J = 8.8 Hz, 2 H, 2 x CHCO), 3.93 (t, J = 6.6 Hz, 2 H, OCH₂), 2.59-2.46 (m, 4 H, CH₂SCH₂), 1.80-1.72 (m, 2 H, OCH₂CH₂), 1.62-

1.52 (m, 2 H, CH₂CH₂CH₂S), 1.48-1.20 (m, 16 H, 8 x CH₂), 0.92-0.80 (m, 2 H, CH₂TMS), 0.23 (s, 9 H, alkyne-TMS), 0.02 (s, 9 H, CH₂TMS); ¹³C NMR (CDCl₃, 101 MHz) δ 160.92 (CO), 135.91 (2 x CHCC), 116.11 (2 x CHCO), 116.00 (CCCTMS), 85.85 (CCTMS), 77.61 (CTMS), 69.77 (OCH₂), 33.54 (CH₂CH₂CH₂S), 31.64 (CH₂CH₂CH₂S), 31.41, 31.33, 31.09, 30.53, 30.72, 30.31, 30.00, 29.61 (8 x CH₂), 29.53 (OCH₂CH₂), 27.79 (SCH₂CH₂TMS), 19.18 (CH₂TMS), 0.02 (C-alkyne-TMS), -0.00 (CH₂TMS); MS (EI⁺) m/z = 490.93 [M]; IR (ATR) 3089.72 (C=C arom), 2944.15 (CH), 2202.34 (alkyne).

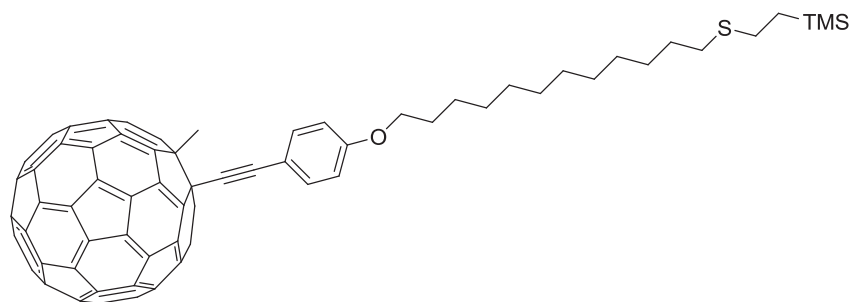
12-(4-ethynyl phenoxy)-dodecyl thioethyltrimethylsilane 7.24



To a solution of **7.23** (4.95 g, 7.92 mmol, 1 eq.) in THF:MeOH (1:1, 400 mL), was added K₂CO₃ (1.6 g, 11.88 mmol, 1.5 eq). The reaction was stirred at r.t. for 2 h, then quenched with H₂O and extracted with DCM (3 x 125 mL). The organic layer was dried over MgSO₄ and evaporated to afford the title compound **7.24** as a yellow solid (3.05 g, 6.21 mmol, 78%). TLC R_f = 0.4 (hexane:DCM, 10:1); ¹H NMR (CDCl₃, 400 MHz) δ 7.39 (d, *J* = 8.8 Hz, 2 H, 2 x CHCC), 6.81 (d, *J* = 8.8 Hz, 2 H, 2 x CHCO), 3.92 (t, *J* = 6.6 Hz, 2 H, OCH₂), 2.98 (s, 1 H, CH-alkyne), 2.65-2.42 (m, 4 H, CH₂SCH₂), 1.84-1.69 (m, 2 H, OCH₂CH₂), 1.64-1.51 (m, 2 H, CH₂CH₂S), 1.48-1.15 (m, 16 H, 8 x CH₂), 0.92-0.79 (m, 2 H, CH₂TMS), 0.01 (s, 9 H, TMS); ¹³C NMR (CDCl₃, 101 MHz) δ 161.22 (CO), 135.24 (2 x CHCC), 116.13

(2 x **CHCO**), 115.58 (**CCCH**), 85.48 (**CCH-alkyne**), 77.42 (**CH-alkyne**), 69.72 (**OCH₂**), 33.62 (**CH₂CH₂CH₂S**), 31.31 (**CH₂CH₂CH₂S**), 31.30, 31.28, 31.26, 30.88, 30.72, 30.22, 30.03, 29.58 (8 x **CH₂**), 29.72 (**OCH₂CH₂**), 27.72 (**SCH₂CH₂TMS**), 19.08 (**CH₂TMS**), -0.00 (**TMS**); **MS** (FI^+) m/z = 490.93 [**M**]; **IR** (ATR) 3290.52 (**CCH-alkyne**), 3087.88 (**C=C arom**), 2943.69 (**CH**), 2207.29 (**alkyne**).

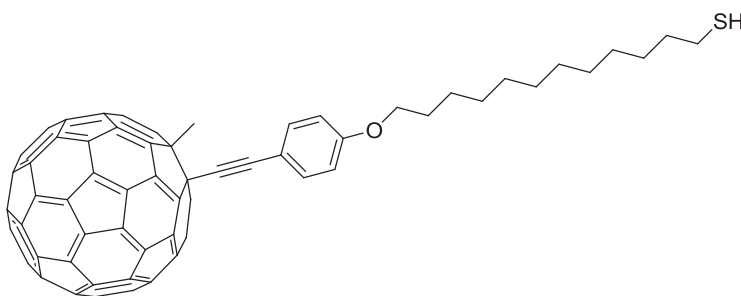
Ethynylphenoxy-4-(dodecyl thioethyltrimethylsilane)-methyl fullerene **7.25**



To a solution of **7.24** (290.5 mg, 0.69 mmol, 1 eq.) in toluene (20 mL) was added LHMDS (1 M in THF, 1.4 mL, 1.4 mmol, 2 eq.). The reaction is stirred at r.t. for 15 min, then a solution of **C₆₀** (1 g, 1.4 mmol, 2 eq.) in toluene (80 mL) was added. After 15 min the reaction was quenched with MeI (2 mL, 13.88 mmol, 20 eq.), and stirred at r.t. for other 3 h. The reaction was evaporated, the product precipitated from toluene with PE, hexane and filtered over silica with toluene to afford the title compound **7.25** as a brown solid (140.8 mg, 0.12 mmol, 18%). **TLC** R_f = 0.95 (toluene:DCM:hexane, 2:2:6); **¹H NMR** (toluene- d_8 , 400 MHz) δ 6.81 (br, 4 H, 4 x **CH-arom**), 3.77 (br, 2 H, **OCH₂**), 2.62-2.38 (m, 4 H, **CH₂SCH₂**), 1.64 (s, 23 H, **OCH₂CH₂**, **CH₂CH₂CH₂S**, **CH₃**, 8 x **CH₂**), 0.90 (br, 2 H, **CH₂TMS**), -0.00

(s, 9 H, TMS); **¹³C NMR** (toluene-d₈, 126 MHz) δ 148.14-116.56 (C-fullerene, CH-arom, C-arom), 70.12 (OCH₂), 53.74 (C-alkyne-C₆₀), 49.87 (C-alkyne-Ar), 34.19 (SCH₂CH₂TMS), 32.20, 32.09, 31.99, 31.95, 31.87, 31.76, 31.60, 31.47 (8 x CH₂), 31.37 (CH₂CH₂CH₂S), 29.82 (CH₂CH₂CH₂S), 28.47 (CH₃), 23.24 (OCH₂CH₂), 19.46 (CH₂TMS), 0.18 (TMS); **MS** (LD⁺) m/z 1153.42 [M]; **IR** (ATR) 3097.93 (C=C arom), 2934.92 (CH) , 2201.72 (alkyne).

Ethynylphenoxy-4-(dodecyl thio)-methyl fullerene **7.26**



To a solution of **7.25** (14 mg, 0.012 mmol, 1 eq.) in toluene (dry, 15 mL) was added TBAF (1 M in THF, 0.5 mL, 0.5 mmol, 40 eq.) and stirred at r.t. for 2 h. The reaction was quenched with MeOH (15 mL), and TFA (100 μL). The reaction mixture was evaporated and purified by precipitation with MeOH to afford the title compound **7.26** as a black solid (9.5 mg, 0.009 mmol, 75%). **TLC** R_f = 0.0 (toluene:MeOH, 8:2); **MS** (LD⁺) m/z 1053.18 [M]; **IR** (ATR) 3091.93 (C=C arom), 2944.17 (CH) , 2202.28 (alkyne).

NMR spectra could not be obtained due to the poor solubility of the compound. After the deprotection the immediate formation of disulphides was observed.

References

- [1] H. Bohme and B. Mencke, *Arch. Pharm.* **1959**, 92, 54-67.
- [2] S. H. Kim, S. H. Lee, and S. H. Kang, *Tet. Lett.* **1998**, 39, 9693-96.
- [3] K. R. Prabhu, A. R. Ramesha, and S. Chandrasekaran, *J. Org. Chem.* **1995**, 60, 7142-43.
- [4] O. Aberkane, J. L. Mieloszynski, D. Robert, M. Born, and D. Paquer, *Phosphorus. Sulfur. Silicon Relat. Elem.* **1993**, 79, 245-56.
- [5] L. Salerno, V. Pittalà, G. Romeo, M. N. Modica, A. Marrazzo, M. A. Siracusa, V. Sorrenti, C. Di Giacomo, L. Vanella, N. N. Parayath, and K. Greish, *Eur. J. Med. Chem.* **2015**, 96, 162-72.
- [6] C. J. Yu, Y. Chong, J. F. Kayyem, and M. Gozin, *J. Org. Chem.* **1999**, 64, 2070-79.
- [7] Y. Shirai, L. Cheng, B. Chen, and J. M. Tour, *J. Am. Chem. Soc.* **2006**, 128, 13479-89.

Summary

The work described in this thesis deals with several aspects of carbohydrates and their biological implications. The library of sugars we synthesised was applied to different macromolecular systems, revealing new chemical and biological features. With two different proteins, the monomeric Np276 and the multimeric Q β , we presented a new strategy that makes different chemical synthetic glycosylation options available to biochemists. We combined three glycosylating strategies - via IME, CuAAC and NCS - in a single precursor that can be easily diversified in the respective glycosylating reagents. Thanks to this triply-divergent approach, different reactions can be tested for protein glycosylation without the need to manipulate the protein and with minimal handling of the sugar counterpart. Our sugar library was conjugated to amino-functionalised magnetic nanoparticles (as tagging agents) to study the biodistribution of this set of carbohydrates *in vivo*. We prepared 20 sugars of different complexity, presenting some of the most relevant derivatisations present in nature or involved in physiological and pathological processes, such as N-Acetylation, fucosylation and sialylation. We tested and compared these sugars on a single platform in animal models, to describe one of their most basic features, the foundation of their interactions: their localisation. Most of the glyconanoparticles accumulated in the liver and spleen, but by studying the minor distribution in other organs we found an unprecedented targeting of muscles by lactose and 1-3,N-Acetylactosamine. The Glycomap obtained represents the first effort in studying the influence of chemical and structural differences in the carbohydrate structures on their biological distribution.

THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE LICENSE TERMS AND CONDITIONS

Jan 24, 2016

This is a License Agreement between Sonia De Munari ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by The American Association for the Advancement of Science, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3795441024406
License date	Jan 24, 2016
Licensed content publisher	The American Association for the Advancement of Science
Licensed content publication	Science
Licensed content title	Intracellular Functions of N-Linked Glycans
Licensed content author	Ari Helenius, and Markus Aebl
Licensed content date	Mar 23, 2001
Volume number	291
Issue number	5512
Type of Use	Thesis / Dissertation
Requestor type	Scientist/individual at a research institution
Format	Print and electronic
Portion	Figure
Number of figures/tables	1
Order reference number	None
Title of your thesis / dissertation	Biological investigation of glycosylated macromolecular structures
Expected completion date	Jan 2016
Estimated size(pages)	500
Total	0.00 GBP

Terms and Conditions

American Association for the Advancement of Science TERMS AND CONDITIONS

Regarding your request, we are pleased to grant you non-exclusive, non-transferable permission, to republish the AAAS material identified above in your work identified above, subject to the terms and conditions herein. We must be contacted for permission for any uses other than those specifically identified in your request above.

The following credit line must be printed along with the AAAS material: "From [Full Reference Citation]. Reprinted with permission from AAAS."

All required credit lines and notices must be visible any time a user accesses any part of the AAAS material and must appear on any printed copies and authorized user might make.

This permission does not apply to figures / photos / artwork or any other content or materials included in your work that are credited to non-AAAS sources. If the requested material is

sourced to or references non-AAAS sources, you must obtain authorization from that source as well before using that material. You agree to hold harmless and indemnify AAAS against any claims arising from your use of any content in your work that is credited to non-AAAS sources.

If the AAAS material covered by this permission was published in Science during the years 1974 - 1994, you must also obtain permission from the author, who may grant or withhold permission, and who may or may not charge a fee if permission is granted. See original article for author's address. This condition does not apply to news articles.

The AAAS material may not be modified or altered except that figures and tables may be modified with permission from the author. Author permission for any such changes must be secured prior to your use.

Whenever possible, we ask that electronic uses of the AAAS material permitted herein include a hyperlink to the original work on AAAS's website (hyperlink may be embedded in the reference citation).

AAAS material reproduced in your work identified herein must not account for more than 30% of the total contents of that work.

AAAS must publish the full paper prior to use of any text.

AAAS material must not imply any endorsement by the American Association for the Advancement of Science.

This permission is not valid for the use of the AAAS and/or Science logos.

AAAS makes no representations or warranties as to the accuracy of any information contained in the AAAS material covered by this permission, including any warranties of merchantability or fitness for a particular purpose.

If permission fees for this use are waived, please note that AAAS reserves the right to charge for reproduction of this material in the future.

Permission is not valid unless payment is received within sixty (60) days of the issuance of this permission. If payment is not received within this time period then all rights granted herein shall be revoked and this permission will be considered null and void.

In the event of breach of any of the terms and conditions herein or any of CCC's Billing and Payment terms and conditions, all rights granted herein shall be revoked and this permission will be considered null and void.

AAAS reserves the right to terminate this permission and all rights granted herein at its discretion, for any purpose, at any time. In the event that AAAS elects to terminate this permission, you will have no further right to publish, publicly perform, publicly display, distribute or otherwise use any matter in which the AAAS content had been included, and all fees paid hereunder shall be fully refunded to you. Notification of termination will be sent to the contact information as supplied by you during the request process and termination shall be immediate upon sending the notice. Neither AAAS nor CCC shall be liable for any costs, expenses, or damages you may incur as a result of the termination of this permission, beyond the refund noted above.

This Permission may not be amended except by written document signed by both parties.

The terms above are applicable to all permissions granted for the use of AAAS material.

Below you will find additional conditions that apply to your particular type of use.

FOR A THESIS OR DISSERTATION

If you are using figure(s)/table(s), permission is granted for use in print and electronic versions of your dissertation or thesis. A full text article may be used in print versions only of a dissertation or thesis.

Permission covers the distribution of your dissertation or thesis on demand by ProQuest / UMI, provided the AAAS material covered by this permission remains in situ.

If you are an Original Author on the AAAS article being reproduced, please refer to your

License to Publish for rules on reproducing your paper in a dissertation or thesis.

FOR JOURNALS:

Permission covers both print and electronic versions of your journal article, however the AAAS material may not be used in any manner other than within the context of your article.

FOR BOOKS/TEXTBOOKS:

If this license is to reuse figures/tables, then permission is granted for non-exclusive world rights in all languages in both print and electronic formats (electronic formats are defined below).

If this license is to reuse a text excerpt or a full text article, then permission is granted for non-exclusive world rights in English only. You have the option of securing either print or electronic rights or both, but electronic rights are not automatically granted and do garner additional fees. Permission for translations of text excerpts or full text articles into other languages must be obtained separately.

Licenses granted for use of AAAS material in electronic format books/textbooks are valid only in cases where the electronic version is equivalent to or substitutes for the print version of the book/textbook. The AAAS material reproduced as permitted herein must remain in situ and must not be exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit), except in the case of permitted textbook companions as noted below.

You must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher."

If your book is an academic textbook, permission covers the following companions to your textbook, provided such companions are distributed only in conjunction with your textbook at no additional cost to the user:

- Password-protected website
- Instructor's image CD/DVD and/or PowerPoint resource
- Student CD/DVD

All companions must contain instructions to users that the AAAS material may be used for non-commercial, classroom purposes only. Any other uses require the prior written permission from AAAS.

If your license is for the use of AAAS Figures/Tables, then the electronic rights granted herein permit use of the Licensed Material in any Custom Databases that you distribute the electronic versions of your textbook through, so long as the Licensed Material remains within the context of a chapter of the title identified in your request and cannot be downloaded by a user as an independent image file.

Rights also extend to copies/files of your Work (as described above) that you are required to provide for use by the visually and/or print disabled in compliance with state and federal laws.

This permission only covers a single edition of your work as identified in your request.

FOR NEWSLETTERS:

Permission covers print and/or electronic versions, provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit)

FOR ANNUAL REPORTS:

Permission covers print and electronic versions provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit)

FOR PROMOTIONAL/MARKETING USES:

Permission covers the use of AAAS material in promotional or marketing pieces such as information packets, media kits, product slide kits, brochures, or flyers limited to a single print run. The AAAS Material may not be used in any manner which implies endorsement or promotion by the American Association for the Advancement of Science (AAAS) or Science of any product or service. AAAS does not permit the reproduction of its name, logo or text on promotional literature.

If permission to use a full text article is permitted, The Science article covered by this permission must not be altered in any way. No additional printing may be set onto an article copy other than the copyright credit line required above. Any alterations must be approved in advance and in writing by AAAS. This includes, but is not limited to, the placement of sponsorship identifiers, trademarks, logos, rubber stamping or self-adhesive stickers onto the article copies.

Additionally, article copies must be a freestanding part of any information package (i.e. media kit) into which they are inserted. They may not be physically attached to anything, such as an advertising insert, or have anything attached to them, such as a sample product. Article copies must be easily removable from any kits or informational packages in which they are used. The only exception is that article copies may be inserted into three-ring binders.

FOR CORPORATE INTERNAL USE:

The AAAS material covered by this permission may not be altered in any way. No additional printing may be set onto an article copy other than the required credit line. Any alterations must be approved in advance and in writing by AAAS. This includes, but is not limited to the placement of sponsorship identifiers, trademarks, logos, rubber stamping or self-adhesive stickers onto article copies.

If you are making article copies, copies are restricted to the number indicated in your request and must be distributed only to internal employees for internal use.

If you are using AAAS Material in Presentation Slides, the required credit line must be visible on the slide where the AAAS material will be reprinted

If you are using AAAS Material on a CD, DVD, Flash Drive, or the World Wide Web, you must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher." Access to any such CD, DVD, Flash Drive or Web page must be restricted to your organization's employees only.

FOR CME COURSE and SCIENTIFIC SOCIETY MEETINGS:

Permission is restricted to the particular Course, Seminar, Conference, or Meeting indicated in your request. If this license covers a text excerpt or a Full Text Article, access to the reprinted AAAS material must be restricted to attendees of your event only (if you have been granted electronic rights for use of a full text article on your website, your website must be password protected, or access restricted so that only attendees can access the content on your site).

If you are using AAAS Material on a CD, DVD, Flash Drive, or the World Wide Web, you must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher."

FOR POLICY REPORTS:

These rights are granted only to non-profit organizations and/or government agencies. Permission covers print and electronic versions of a report, provided the required credit line appears in both versions and provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately.

FOR CLASSROOM PHOTOCOPIES:

Permission covers distribution in print copy format only. Article copies must be freestanding and not part of a course pack. They may not be physically attached to anything or have anything attached to them.

FOR COURSEPACKS OR COURSE WEBSITES:

These rights cover use of the AAAS material in one class at one institution. Permission is valid only for a single semester after which the AAAS material must be removed from the Electronic Course website, unless new permission is obtained for an additional semester. If the material is to be distributed online, access must be restricted to students and instructors enrolled in that particular course by some means of password or access control.

FOR WEBSITES:

You must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher."

Permissions for the use of Full Text articles on third party websites are granted on a case by case basis and only in cases where access to the AAAS Material is restricted by some means of password or access control. Alternately, an E-Print may be purchased through our reprints department (brocheleau@rockwaterinc.com).

REGARDING FULL TEXT ARTICLE USE ON THE WORLD WIDE WEB IF YOU ARE AN 'ORIGINAL AUTHOR' OF A SCIENCE PAPER

If you chose "Original Author" as the Requestor Type, you are warranting that you are one of authors listed on the License Agreement as a "Licensed content author" or that you are acting on that author's behalf to use the Licensed content in a new work that one of the authors listed on the License Agreement as a "Licensed content author" has written.

Original Authors may post the 'Accepted Version' of their full text article on their personal or on their University website and not on any other website. The 'Accepted Version' is the version of the paper accepted for publication by AAAS including changes resulting from peer review but prior to AAAS's copy editing and production (in other words not the AAAS published version).

FOR MOVIES / FILM / TELEVISION:

Permission is granted to use, record, film, photograph, and/or tape the AAAS material in connection with your program/film and in any medium your program/film may be shown or heard, including but not limited to broadcast and cable television, radio, print, world wide

web, and videocassette.

The required credit line should run in the program/film's end credits.

FOR MUSEUM EXHIBITIONS:

Permission is granted to use the AAAS material as part of a single exhibition for the duration of that exhibit. Permission for use of the material in promotional materials for the exhibit must be cleared separately with AAAS (please contact us at permissions@aaas.org).

FOR TRANSLATIONS:

Translation rights apply only to the language identified in your request summary above.

The following disclaimer must appear with your translation, on the first page of the article, after the credit line: "This translation is not an official translation by AAAS staff, nor is it endorsed by AAAS as accurate. In crucial matters, please refer to the official English-language version originally published by AAAS."

FOR USE ON A COVER:

Permission is granted to use the AAAS material on the cover of a journal issue, newsletter issue, book, textbook, or annual report in print and electronic formats provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately.

By using the AAAS Material identified in your request, you agree to abide by all the terms and conditions herein.

Questions about these terms can be directed to the AAAS Permissions department

permissions@aaas.org.

Other Terms and Conditions:

v 2

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

**ELSEVIER LICENSE
TERMS AND CONDITIONS**

Jan 24, 2016

This is a License Agreement between Sonia De Munari ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard, Langford Lane Kidlington, Oxford, OX5 1GB, UK
Registered Company Number	1982084
Customer name	Sonia De Munari
Customer address	Chemistry Research Laboratory Oxford, OX13TA
License number	3795421316235
License date	Jan 24, 2016
Licensed content publisher	Elsevier
Licensed content publication	Advanced Drug Delivery Reviews
Licensed content title	Endogenous lectins as targets for drug delivery
Licensed content author	N. Yamazaki, S. Kojima, N.V. Bovin, S. André, S. Gabius, H.-J. Gabius
Licensed content date	30 September 2000
Licensed content volume number	43
Licensed content issue number	2-3
Number of pages	20
Start Page	225
End Page	244
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables /illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	12
Title of your thesis/dissertation	Biological investigation of glycosylated macromolecular structures
Expected completion date	Jan 2016

Estimated size (number of pages)	500
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 GBP
VAT/Local Sales Tax	0.00 GBP / 0.00 GBP
Total	0.00 GBP

[Terms and Conditions](#)

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never

granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation:** This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com>; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at <http://www.elsevier.com>. All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. For journal authors: the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
 - o via their non-commercial person homepage or blog
 - o by updating a preprint in arXiv or RePEc with the accepted manuscript
 - o via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
 - o directly by providing copies to their students or to research collaborators for their personal use
 - o for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- after the embargo period
 - o via non-commercial hosting platforms such as their institutional repository
 - o via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a [CrossMark logo](#), the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's [posting policy](#) for further information.

18. **For book authors** the following clauses are applicable in addition to the above:

Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation:** If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our [open access license policy](#) for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are

available at <http://creativecommons.org/licenses/by/4.0>.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-nd/4.0>. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. Other Conditions:

v1.8

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

**JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS**

Jan 24, 2016

This Agreement between Sonia De Munari ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	3795441395776
License date	Jan 24, 2016
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Proteomics
Licensed Content Title	Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds
Licensed Content Author	David J. Harvey,Anthony H. Merry,Louise Royle,Matthew P. Campbell,Raymond A. Dwek,Pauline M. Rudd
Licensed Content Date	Aug 7, 2009
Pages	6
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Table 1
Will you be translating?	No
Title of your thesis / dissertation	Biological investigation of glycosylated macromolecular structures
Expected completion date	Jan 2016
Expected size (number of pages)	500
Requestor Location	Sonia De Munari Chemistry Research Laboratory 12 Mansfield Road Oxford, United Kingdom OX13TA Attn: Sonia De Munari
Billing Type	Invoice
Billing Address	Sonia De Munari Chemistry Research Laboratory 12 Mansfield Road Oxford, United Kingdom OX13TA Attn: Sonia De Munari
Total	0.00 GBP
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at <http://myaccount.copyright.com>).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.
- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, **and any CONTENT (PDF or image file) purchased as part of your order**, is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.
- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner. **For STM Signatory Publishers clearing permission under the terms of the [STM Permissions Guidelines](#) only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts**, You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.
- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the

continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.
- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.
- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.
- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.
- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.
- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or

excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.
- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.
- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.
- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.
- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.
- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The [Creative Commons Attribution License \(CC-BY\)](#) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The [Creative Commons Attribution Non-Commercial \(CC-BY-NC\) License](#) permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The [Creative Commons Attribution Non-Commercial-NoDerivs License](#) (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library <http://olabout.wiley.com/WileyCDA/Section/id-410895.html>

Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS

Jan 24, 2016

This is a License Agreement between Sonia De Munari ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3795450103264
License date	Jan 24, 2016
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature Materials
Licensed content title	The rise of graphene
Licensed content author	A. K. Geim and K. S. Novoselov
Licensed content date	Mar 1, 2007
Volume number	6
Issue number	3
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables /illustrations	1
High-res required	no
Figures	Fig. 1
Author of this NPG article	no
Your reference number	None
Title of your thesis / dissertation	Biological investigation of glycosylated macromolecular structures
Expected completion date	Jan 2016
Estimated size (number of pages)	500
Total	0.00 GBP

Terms and Conditions

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use, subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested indicates that it was reprinted or adapted by NPG with permission from another source,

then you should also seek permission from that source to reuse the material.

2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version. Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run). NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.
3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).
4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.
5. The credit line should read:
Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
For AOP papers, the credit line should read:
Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

Note: For republication from the *British Journal of Cancer*, the following credit lines apply.

Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)
For AOP papers, the credit line should read:
Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)

Note: For adaptation from the *British Journal of Cancer*, the following credit line applies.

Adapted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit <http://www.macmillanmedicalcommunications.com> for more information. Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows:

Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication).

Note: For translation from the *British Journal of Cancer*, the following credit line applies.

Translated by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you.

Special Terms:

v1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE LICENSE TERMS AND CONDITIONS

Jan 24, 2016

This is a License Agreement between Sonia De Munari ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by The American Association for the Advancement of Science, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3795450464658
License date	Jan 24, 2016
Licensed content publisher	The American Association for the Advancement of Science
Licensed content publication	Science
Licensed content title	Automated Solid-Phase Synthesis of Oligosaccharides
Licensed content author	Obadiah J. Plante, Emma R. Palmacci, Peter H. Seeberger
Licensed content date	Feb 23, 2001
Volume number	291
Issue number	5508
Type of Use	Thesis / Dissertation
Requestor type	Scientist/individual at a research institution
Format	Print and electronic
Portion	Figure
Number of figures/tables	1
Order reference number	None
Title of your thesis / dissertation	Biological investigation of glycosylated macromolecular structures
Expected completion date	Jan 2016
Estimated size(pages)	500
Total	0.00 GBP

Terms and Conditions

American Association for the Advancement of Science TERMS AND CONDITIONS

Regarding your request, we are pleased to grant you non-exclusive, non-transferable permission, to republish the AAAS material identified above in your work identified above, subject to the terms and conditions herein. We must be contacted for permission for any uses other than those specifically identified in your request above.

The following credit line must be printed along with the AAAS material: "From [Full Reference Citation]. Reprinted with permission from AAAS."

All required credit lines and notices must be visible any time a user accesses any part of the AAAS material and must appear on any printed copies and authorized user might make.

This permission does not apply to figures / photos / artwork or any other content or materials included in your work that are credited to non-AAAS sources. If the requested material is

sourced to or references non-AAAS sources, you must obtain authorization from that source as well before using that material. You agree to hold harmless and indemnify AAAS against any claims arising from your use of any content in your work that is credited to non-AAAS sources.

If the AAAS material covered by this permission was published in Science during the years 1974 - 1994, you must also obtain permission from the author, who may grant or withhold permission, and who may or may not charge a fee if permission is granted. See original article for author's address. This condition does not apply to news articles.

The AAAS material may not be modified or altered except that figures and tables may be modified with permission from the author. Author permission for any such changes must be secured prior to your use.

Whenever possible, we ask that electronic uses of the AAAS material permitted herein include a hyperlink to the original work on AAAS's website (hyperlink may be embedded in the reference citation).

AAAS material reproduced in your work identified herein must not account for more than 30% of the total contents of that work.

AAAS must publish the full paper prior to use of any text.

AAAS material must not imply any endorsement by the American Association for the Advancement of Science.

This permission is not valid for the use of the AAAS and/or Science logos.

AAAS makes no representations or warranties as to the accuracy of any information contained in the AAAS material covered by this permission, including any warranties of merchantability or fitness for a particular purpose.

If permission fees for this use are waived, please note that AAAS reserves the right to charge for reproduction of this material in the future.

Permission is not valid unless payment is received within sixty (60) days of the issuance of this permission. If payment is not received within this time period then all rights granted herein shall be revoked and this permission will be considered null and void.

In the event of breach of any of the terms and conditions herein or any of CCC's Billing and Payment terms and conditions, all rights granted herein shall be revoked and this permission will be considered null and void.

AAAS reserves the right to terminate this permission and all rights granted herein at its discretion, for any purpose, at any time. In the event that AAAS elects to terminate this permission, you will have no further right to publish, publicly perform, publicly display, distribute or otherwise use any matter in which the AAAS content had been included, and all fees paid hereunder shall be fully refunded to you. Notification of termination will be sent to the contact information as supplied by you during the request process and termination shall be immediate upon sending the notice. Neither AAAS nor CCC shall be liable for any costs, expenses, or damages you may incur as a result of the termination of this permission, beyond the refund noted above.

This Permission may not be amended except by written document signed by both parties.

The terms above are applicable to all permissions granted for the use of AAAS material.

Below you will find additional conditions that apply to your particular type of use.

FOR A THESIS OR DISSERTATION

If you are using figure(s)/table(s), permission is granted for use in print and electronic versions of your dissertation or thesis. A full text article may be used in print versions only of a dissertation or thesis.

Permission covers the distribution of your dissertation or thesis on demand by ProQuest / UMI, provided the AAAS material covered by this permission remains in situ.

If you are an Original Author on the AAAS article being reproduced, please refer to your

License to Publish for rules on reproducing your paper in a dissertation or thesis.

FOR JOURNALS:

Permission covers both print and electronic versions of your journal article, however the AAAS material may not be used in any manner other than within the context of your article.

FOR BOOKS/TEXTBOOKS:

If this license is to reuse figures/tables, then permission is granted for non-exclusive world rights in all languages in both print and electronic formats (electronic formats are defined below).

If this license is to reuse a text excerpt or a full text article, then permission is granted for non-exclusive world rights in English only. You have the option of securing either print or electronic rights or both, but electronic rights are not automatically granted and do garner additional fees. Permission for translations of text excerpts or full text articles into other languages must be obtained separately.

Licenses granted for use of AAAS material in electronic format books/textbooks are valid only in cases where the electronic version is equivalent to or substitutes for the print version of the book/textbook. The AAAS material reproduced as permitted herein must remain in situ and must not be exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit), except in the case of permitted textbook companions as noted below.

You must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher."

If your book is an academic textbook, permission covers the following companions to your textbook, provided such companions are distributed only in conjunction with your textbook at no additional cost to the user:

- Password-protected website
- Instructor's image CD/DVD and/or PowerPoint resource
- Student CD/DVD

All companions must contain instructions to users that the AAAS material may be used for non-commercial, classroom purposes only. Any other uses require the prior written permission from AAAS.

If your license is for the use of AAAS Figures/Tables, then the electronic rights granted herein permit use of the Licensed Material in any Custom Databases that you distribute the electronic versions of your textbook through, so long as the Licensed Material remains within the context of a chapter of the title identified in your request and cannot be downloaded by a user as an independent image file.

Rights also extend to copies/files of your Work (as described above) that you are required to provide for use by the visually and/or print disabled in compliance with state and federal laws.

This permission only covers a single edition of your work as identified in your request.

FOR NEWSLETTERS:

Permission covers print and/or electronic versions, provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit)

FOR ANNUAL REPORTS:

Permission covers print and electronic versions provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit)

FOR PROMOTIONAL/MARKETING USES:

Permission covers the use of AAAS material in promotional or marketing pieces such as information packets, media kits, product slide kits, brochures, or flyers limited to a single print run. The AAAS Material may not be used in any manner which implies endorsement or promotion by the American Association for the Advancement of Science (AAAS) or Science of any product or service. AAAS does not permit the reproduction of its name, logo or text on promotional literature.

If permission to use a full text article is permitted, The Science article covered by this permission must not be altered in any way. No additional printing may be set onto an article copy other than the copyright credit line required above. Any alterations must be approved in advance and in writing by AAAS. This includes, but is not limited to, the placement of sponsorship identifiers, trademarks, logos, rubber stamping or self-adhesive stickers onto the article copies.

Additionally, article copies must be a freestanding part of any information package (i.e. media kit) into which they are inserted. They may not be physically attached to anything, such as an advertising insert, or have anything attached to them, such as a sample product. Article copies must be easily removable from any kits or informational packages in which they are used. The only exception is that article copies may be inserted into three-ring binders.

FOR CORPORATE INTERNAL USE:

The AAAS material covered by this permission may not be altered in any way. No additional printing may be set onto an article copy other than the required credit line. Any alterations must be approved in advance and in writing by AAAS. This includes, but is not limited to the placement of sponsorship identifiers, trademarks, logos, rubber stamping or self-adhesive stickers onto article copies.

If you are making article copies, copies are restricted to the number indicated in your request and must be distributed only to internal employees for internal use.

If you are using AAAS Material in Presentation Slides, the required credit line must be visible on the slide where the AAAS material will be reprinted

If you are using AAAS Material on a CD, DVD, Flash Drive, or the World Wide Web, you must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher." Access to any such CD, DVD, Flash Drive or Web page must be restricted to your organization's employees only.

FOR CME COURSE and SCIENTIFIC SOCIETY MEETINGS:

Permission is restricted to the particular Course, Seminar, Conference, or Meeting indicated in your request. If this license covers a text excerpt or a Full Text Article, access to the reprinted AAAS material must be restricted to attendees of your event only (if you have been granted electronic rights for use of a full text article on your website, your website must be password protected, or access restricted so that only attendees can access the content on your site).

If you are using AAAS Material on a CD, DVD, Flash Drive, or the World Wide Web, you must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher."

FOR POLICY REPORTS:

These rights are granted only to non-profit organizations and/or government agencies. Permission covers print and electronic versions of a report, provided the required credit line appears in both versions and provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately.

FOR CLASSROOM PHOTOCOPIES:

Permission covers distribution in print copy format only. Article copies must be freestanding and not part of a course pack. They may not be physically attached to anything or have anything attached to them.

FOR COURSEPACKS OR COURSE WEBSITES:

These rights cover use of the AAAS material in one class at one institution. Permission is valid only for a single semester after which the AAAS material must be removed from the Electronic Course website, unless new permission is obtained for an additional semester. If the material is to be distributed online, access must be restricted to students and instructors enrolled in that particular course by some means of password or access control.

FOR WEBSITES:

You must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher."

Permissions for the use of Full Text articles on third party websites are granted on a case by case basis and only in cases where access to the AAAS Material is restricted by some means of password or access control. Alternately, an E-Print may be purchased through our reprints department (brocheleau@rockwaterinc.com).

REGARDING FULL TEXT ARTICLE USE ON THE WORLD WIDE WEB IF YOU ARE AN 'ORIGINAL AUTHOR' OF A SCIENCE PAPER

If you chose "Original Author" as the Requestor Type, you are warranting that you are one of authors listed on the License Agreement as a "Licensed content author" or that you are acting on that author's behalf to use the Licensed content in a new work that one of the authors listed on the License Agreement as a "Licensed content author" has written.

Original Authors may post the 'Accepted Version' of their full text article on their personal or on their University website and not on any other website. The 'Accepted Version' is the version of the paper accepted for publication by AAAS including changes resulting from peer review but prior to AAAS's copy editing and production (in other words not the AAAS published version).

FOR MOVIES / FILM / TELEVISION:

Permission is granted to use, record, film, photograph, and/or tape the AAAS material in connection with your program/film and in any medium your program/film may be shown or heard, including but not limited to broadcast and cable television, radio, print, world wide

web, and videocassette.

The required credit line should run in the program/film's end credits.

FOR MUSEUM EXHIBITIONS:

Permission is granted to use the AAAS material as part of a single exhibition for the duration of that exhibit. Permission for use of the material in promotional materials for the exhibit must be cleared separately with AAAS (please contact us at permissions@aaas.org).

FOR TRANSLATIONS:

Translation rights apply only to the language identified in your request summary above.

The following disclaimer must appear with your translation, on the first page of the article, after the credit line: "This translation is not an official translation by AAAS staff, nor is it endorsed by AAAS as accurate. In crucial matters, please refer to the official English-language version originally published by AAAS."

FOR USE ON A COVER:

Permission is granted to use the AAAS material on the cover of a journal issue, newsletter issue, book, textbook, or annual report in print and electronic formats provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately.

By using the AAAS Material identified in your request, you agree to abide by all the terms and conditions herein.

Questions about these terms can be directed to the AAAS Permissions department

permissions@aaas.org.

Other Terms and Conditions:

v 2

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.