UNIVERSITA' DEGLI STUDI DI MILANO-BICOCCA

Facoltà di Scienze MM. FF. NN

Corso di Dottorato in Scienze Chimiche-XXI ciclo

Glycomimetics: design synthesis and biological activity studies

Luca Gabrielli

UNIVERSITA' DEGLI STUDI DI MILANO-BICOCCA
Facoltà di Scienze MM. FF. NN
Corso di Dottorato in Scienze Chimiche-XXI ciclo

Glycomimetics: design synthesis and biological activity studies

PhD Student: Luca Gabrielli

Supervisor: Prof. Laura Cipolla

PhD Coordinator: Prof. Roberto Todeschini

Milan, 2012

“Well, I'll eat it, and if it makes me grow larger, I can reach the key; and if it makes me grow smaller, I can creep under the door; so either way I'll get into the garden, and I don't care which happens!”

CONTENTS

1. Background: Nature's Sweet Code

A substantial part of organic matter on earth is constituted by carbohydrates mostly in the form of cellulose from plants and chitin from arthropods and fungi. In addition, carbohydrates are found in all known living organisms free or covalently linked to other structures, mainly proteins and lipids by glycosylation, and are usually referred to as glycans[1].

Some of these complex molecules are exposed on the surface of cells or are secreted into the surrounding environment, mediating a wide variety of processes including structural support, protection, recognition, localization, and information/nutrient transfer.

Carbohydrates making up the glycan repertoire of each species can differ dramatically[2,3]. Even if some monosaccharides used for the construction of the glycan are fairly common in metazoan animals, many members of the Bacteria and Archaea domains synthesize a number of species-specific carbohydrates, and some of them are still unknown[4]. The most common classes of monosaccharides include sialic acids (e.g. N-acetylneuraminic acid, N-glycolylneuraminic acid), hexoses (e.g. glucose, mannose, galactose), hexosamines (N-acetylglucosamine, N-acetylgalactosamine), deoxyhexoses (e.g. fucose), pentoses (ribose, xylose), and uronic acids (i.e. iduronic acid, glucuronic acid).

The level of diversity introduced in the oligosaccharide structures is extremely high; it is due to the occurrence of the two possible

anomers per glycosidic bond and by the presence of five commonly found potential hydroxyl acceptors per sugar unit, affording highly complex linear or branched structures. Thus in theory, there is virtually no limit to the number of different glycans that can be synthesised by nature. Differently from metazoan animals that are used to synthesise only a limited range of glycan structures, vertebrates, and mammals in particular, possess a highly complex glycan stock that is structurally diverse from that of invertebrates, with lower eukaryotes and prokaryotes.

Glycan structures are synthesised and manipulated by a set of specific enzymes generally referred to as carbohydrate-processing enzymes, which include glycosyltransferases that catalyse the transfer of a sugar from a nucleotide sugar donor or a lipid sugar to an acceptor molecule, and glycosidases, enzymes that catalyse the hydrolysis of glycosidic bonds into glycan structures. Enzymes responsible for carbohydrate processing, are encoded by genes representing >1% of the total genome, and their expression is highly regulated depending on cell metabolism, activation, microenvironmental changes and aging. Therefore, glycosylation itself can give different structural variations to a given protein, leading to identical proteins linked to different glycan structures (glycoforms), giving thousands of potential glycan structures, which adds to the diversity already created by the proteome. The glycans can be considered as the "*sugar code*" of a cell, and the population of sugars in glycoforms governs physiological and

pathological states in mammals[5,6,7]. It is increasingly clear that the glycan structural variation among organisms is a molecular basis for interspecies recognition systems[1]. In addition, glycosylation abnormalities, in mice as well as in humans, linked to disease have shown that the glycome contains a significant amount of biological information[8,9,10]. For example, the impact of glycosylation on the regulation of innate and adaptive immune responses (glycoimmunology) is relatively new. An increasing number of studies reveal the relevance of glycosylation to pathogen recognition, to the modulation of the innate immune system and to the control of immune cell homeostasis and inflammation[11]. Similarly important is the effect of glycan-containing 'information' in the development of autoimmune diseases and cancer[12,13,14]. In addition, glycans from bacteria or other nonvertebrate organisms can modulate the development and activation of the mammalian immune system[15,16,17,18,19,] as well as carbohydrate/protein interactions have long been known to play an important role in viral invasion[20].

Despite the increasing knowledge about natural glycan diversity, flanked by the emergence of families of proteins with carbohydrate-binding activities, assignments of information content for the glycocode will continue to pose a challenge. However, the rapid development of glycomic methods (see for example refs [21,22,23,24,25,26,27] highlighting the rapid development in the field) will greatly facilitate not only the study of proteins for carbohydrate-binding activities, but also the revelation of their ligands. These

approaches will take advantages from advances in structural information gained by mass spectroscopy[28] and NMR techniques[29]. The structural details of carbohydrate–protein complexes can be examined by NMR, providing site-specific information on the architecture, binding selectivity and plasticity of the carbohydrate-binding sites of the proteins. New insights into the conformational behaviour of free and protein-bound glycans[30] pave the way for the design of carbohydrate-based therapeutics.

The potentialities of carbohydrate-based therapeutics are therefore wide [31] , and will expand as far as the glycocode can be deciphered. Different "families" of carbohydrate-based therapeutics can be designed, taking advantage of different potentialities offered by this versatile class of biomolecules. Hence, we can design carbohydrate-based therapeutics following four different approaches, giving rise to the different classes of molecules[32].

1) Native glycan structures, possessing the desired biological activity. This approach suffers from synthetic limitation and metabolic instability of the potential drug.
2) Carbohydrate scaffolds: this approach takes advantage of an unparalleled opportunity to generate libraries of high functional and structural diversity from carbohydrates.
3) Glycomimetics in which proper modifications allow a stronger interaction with the target receptor; this approach should guarantee a higher in vivo stability and better pharmacokinetic properties.

1. Native glycan structures for therapeutics design

Clinical applicability of carbohydrate-based therapeutics was born with the discovery of the human blood groups, although evidence that these were glycan antigens came later[33]. The clinically serious consequences of mismatched blood transfusions resulting from the natural immune responses to the A and B antigens in those who lack them served to focus research onto the chemical structures and the genetic basis of the blood group antigens. The pioneering biochemical and immunochemical work by Morgan's and Watkins' groups in the UK and by Kabat in the USA[34,35,36] elucidated the oligosaccharide sequences that constitute the major blood group antigens. Their work also laid the foundations for the recent research on the roles of oligosaccharides related to these antigens as ligands for effector proteins that have key roles in the leukocytes trafficking in inflammation[37].

The antithrombotic glycan heparin is one of the most commonly used drugs[38,39]; heparin from porcine mucosal tissue has been used as an anticoagulant drug for more than eighty years. However, the need for synthetic heparin for anticoagulation therapy was recently raised by the discovery of heparin batches associated with anaphylactoid-type reactions, causing hypotension and which resulted in almost 100 deaths. These side-reactions were ascribed to contamination with a semi-synthetic over-sulfated chondroitin sulfate, which is a popular shellfish-derived oral supplement for arthritis.

Native glycan structures have been used also for the design of vaccines: specific polysaccharides located on the surface of bacteria or viruses can be specifically recognised by the host immune system. Hence, immunization with polysaccharide antigens should afford protection against pathogens possessing these sugar moieties on their surface [40, 41]. Indeed, CPS and LPS can be recognized by receptors of the innate immune system leading to the production of cytokines, chemokines and cellular adhesion molecules. Glycan-based vaccines have a long history dating back to 1930 when induction of an antibody response in patients by pneumococcus-specific polysaccharides that was later shown to be protective was observed [42, 43]. Since then several capsular polysaccharides have been used for the design of vaccines. With efficient synthetic methods, pure and well-defined carbohydrates can be obtained in sufficient amounts for the construction of vaccines.

Synthetic oligosaccharides have also been employed for the development of therapeutic vaccines for cancer. Cancer cells often display alterations in their glycan repertoire if compared to normal cells, resulting in the accumulation of new structures such as Globo-H, sialyl-Lex (sLex), Ley, sLea, sTn, TF, GM3, Gb3, GM2 on the cell surface (Fig. 2) [44, 45], usually referred to as tumour-associated carbohydrate antigens (TACA). TACAs can be ideal candidates to elicit an antitumor immune response, however to date, no carbohydrate-based anticancer vaccines succeeded to get definitely into the clinic.

However, the development of therapeutics based on native glycan structures has been hampered by synthetic limitations. Unlike proteins and nucleic acids, oligosaccharides are difficult to synthesize chemically, and general methods are lacking for the routine preparation of these important compounds, usually requiring multiple selective protection and deprotection steps; hence, even if recent developments are beginning to address these problems[46,47,48,49,50,51], the chemical synthesis of oligosaccharides is a major undertaking.

2. Carbohydrates as scaffolds for therapeutics design

Carbohydrates present many advantages for the construction of chiral scaffolds, cores or templates, over other biomolecules.

The low cost, abundance, content of the highest density of functional groups of any naturally occurring material and ease with which they can be obtained in a pure state are among the most important features that make carbohydrates prime candidates for the synthesis of chiral scaffolds [52], useful in pharmaceutical and medicinal chemistry[53]. The key issue in this field is the development of better scaffolding structures for the presentation of binding functional groups in a specific orientation. The benefits of such carbohydrate scaffolds include an attractive balance between rigidity and diversity of functional group orientation. For example, in the area of protein-protein interactions, it is generally believed that the size and the rigidity of the structural elements within a small molecule, such as monosaccharides, tend to provide better binding/interfering agents.

The use of carbohydrates as scaffolds for the generation of bioactive compounds was introduced by Hirschmann, Nicolau and Amos Smith III in 1992 for the synthesis of a peptidomimetic directed to the somatostatin receptors, starting from a β-D-glucoside[54]. Since then, the field has rapidly developed, only partially covered in some reviews[55,56,57,58,59]; probably the largest use of carbohydrate scaffolds applies to the synthesis of peptidomimetics[60,61]. A very recent approach in the design of carbohydrate-based drugs involves the engineering of the drug structure onto a sugar moiety; in this case the glycidic part of the molecule might address solubility issues (one of the main problems of organic-based drugs) and/or modulate the pharmacokinetic properties. This idea is very innovative, and the potentialities still have to be deeply investigated. The glycidic moiety of the drug cannot only modulate the solubility and pharmacokinetics, taking into account that each hydroxyl group can be free or eventually acylated or alkylated, but can also target the drug to cells overexpressing specific sugar receptors or enzymes. A recent example is given by the synthesis of $GABA_A$ receptor ligands, where the drug moiety, g-aminobutiric acid and its lactones and lactames[62], or pyrrolobenzodiazepines[63], have been engineered on a fructose scaffold. In this case the sugar moiety plays a crucial role in improving the drug solubility; in addition, it confers conformational constrains to the drug, which is a key issue in the design of $GABA_A$ receptor ligands.

3. Glycomimetics

The understanding of carbohydrate-protein interactions has enabled the development of a new class of small-molecule drugs designed after their structure, known as glycomimetics. Hence, since carbohydrates are an easily accessible starting material, they offer the possibility to synthesise new therapeutics. These compounds mimic native carbohydrate function (the native carbohydrate function or native carbohydrate functions), but reducing the drawbacks of carbohydrate leads, such as their low activity and scarse drug-like properties[64,65,66,67,68].

The use of glycomimetics can offer several advantages over their natural counterparts:

- They usually possess increased metabolic and chemical stability;
- They can be tailor-made with suitable functional groups (pharmacophores) improving the interaction with receptors/enzymes;
- The chemical synthesis can sometimes be simplified by the introduction of functional groups of choice for chemoselective ligation methods, allowing easy conjugation to other chemical entities, i.e. sugars, proteins/peptides, solid support, nanoparticles, dendrimers…

Several modifications of the sugar backbone can be made towards glycomimetics: the endocyclic oxygen can be substitute by a carbon atom (cyclitols or carbasugars) [69] , by a nitrogen atom

(iminosugars)[70], or either by a sulphur atom (thiosugars) or a phosphorus atom (phosphasugars). Alternatively, the exocylic oxygen can be substituted by the same set of atoms, giving respectively C-glycosides, N-glycosides, thioglycosides and P-glycosides.

In some cases glycomimetics are modified in different positions, rather then the endocyclic or exocylic oxygen, through the introduction of pharmacophoric groups on the sugar backbones; in addition, they can also be designed in order to mimic the enzymatic transition state.

Glycomimetics of oligosaccharides, such as Sialyl Lewis X have been proposed as anti-inflammatory agents[71]. sLex tetrasaccharide is the ligand of a class of calcium dependent lectins, including E-, P-, and L-selectins. The selectin-carbohydrate interaction is involved in the inflammatory cascade and, in certain cases, metastasis. A new strategy for the treatment of acute and chronic inflammatory diseases and postsurgical reperfusion injury focuses on the antagonism of lectin-sLex interaction. Lipid A antagonists for sepsis treatment have been designed on the basis of lipid A structure from Rhodobacter capsulatus[72,73]. Among these compounds E5564 (also called Eritoran)[74], is noteworthy since it has shown to be promising in blocking endotoxin challenge; it is now in Phase III clinical trials as an anti-sepsis drug.

Many glycomimetics mimic a monosaccaride structure, such as the anti flu drugs Zanamivir[75] (licensed by GlaxoSmithKline as

Relenza®) and Oseltamivir[76] (licensed by Roche as Tamiflu®), both inhibitors of the viral neuraminidase, that were designed after the structure of monosaccharide sialic acid, and in particular mimick its structure assumed in the transition state of the enzymatic reaction.

2. Aim of the work.

My PhD project has been focused on the design, synthesis and biological activity studies of glicomimetics. During these three years I have been mainly interested in three different arguments that I will elucidate in detail in this thesis; for the sake of simplicity I have decided to name them with the name of the targeted proteins.

1. Arabinose-5-phosphate isomerase (API)

API is an isomerase present in gram-negative bacteria and is a key enzyme implicated in the biosynthesis of the lipopolysaccharides (LPS's). It is responsible for the first step of the biosynthetic pathway of Kdo and catalyses the reversible isomerization from ribulose-5-phosphate (Ru5P) to arabinose-5-phosphate (A5P). I have focused on the design and synthesis of arabinose-based glycomimetics, in order to inhibit API, thus acting as potential antibiotic agents.

2. Protein kinase B (Akt)

Akt is a human protein kinase that phosphorylates diverse protein substrates to promote different cellular responses. For example, many lines of evidence demonstrate that Akt is a critical player in the development, growth, and therapeutic resistance of cancers. Up-regulation and increased Akt activity induce oncogenic transformation of cells and tumour formation in the breast, prostate, ovary and pancreas. Blockage of Akt signalling then results in apoptosis and growth inhibition of tumour cells with elevated Akt activity. In order to synthesise potential anti-cancer agents, I have

worked on the design and synthesis of glucose based inhibitors of Akt.

3. Cation independent mannose-6-phosphate receptor (CI-M6PR)

Mannose-6-phospate (M6P) acts as a trafficking marker for the delivery of enzymes to the lysosome. According to the type of ligand, CI-M6PR may modulate a large panel of biological pathways, such as cell migration, wound healing, angiogenesis, cell growth inhibition, apoptosis or viral infection. Therefore, CI-M6PR constitutes a target for several clinical applications, which justifies the development of synthetic M6P analogues to compensate the poor stability of natural M6P in human serum and to increase the binding stability with the receptor. In order to address these challenges I have been working on the design and synthesis of a new fluorinated analogue of M6P and on a new glycoprotein-molecular ruler for CI-M6P. This work was performed at the B.G. Davis group, Organic Chemistry Department, University of Oxford, Oxford (England).

Even if the target proteins (enzymes and receptors) are very different, these three arguments have a common element. In all cases, starting from the known mechanistic, structural and biological information of the single protein, I have designed glycomimetics that are potentially able to interact with the target protein. All the studied proteins are biologically relevant, so the final aim of the synthetic effort is always a biological test, in order to check the hypothesis done during the

design step and to get new useful information about the biological system. The synthesized compounds are always glycomimetics or, more generally speaking, glycans: they can be arabinose-based mimetic of A5P, glucose-based mimetic of an inisitol structure, or synthetic glycoproteins.

The most used analytical instrument for the characterization of all the synthesized compounds and for the study of their interaction with enzymes is NMR spectroscopy.

The wide use of this analytical technique led me to study topics that were not related to carbohydrates themselves, but related to NMR techniques. Therefore, I increased my knowledge and applied NMR spectroscopy to the biomaterial field and in particular I have studied the reactivity of reaction systems widely used for biomaterial production using NMR. I will briefly discuss these works in the last part of my thesis.

3.1. API: Introduction

Despite advances in the treatment of infectious diseases, pathogenic microorganisms are still one of the most important threats to human (and animal) health worldwide[77]. The fact that many infectious disease agents have never been under control or have remerged, combined with the alarming spread of antibiotic resistance, suggests that investigations into the molecular mechanisms underlying bacterial infections and host interactions deserve strong attention by researchers. The surface of bacterial pathogens is the first site of host interaction and is a major target for the antibacterial activity of the host. Among microbial components, lipopolysaccharides (LPS) in the outer membrane (OM) of Gram-negative bacteria are one key structure that is sensed by the host and represents a potent stimulant of immune responses. LPSs are very well characterized for their immunological, pharmacological and pathophysiological effects displayed in eukaryotic cells and organisms. In general, these amphiphilic molecules comprise three regions, which can be differentiated on the basis of their structures, functions, genetics and biosynthesis: lipid A, the *core* region, and a polysaccharide portion, which may be the *O*-specific polysaccharide, the Enterobacterial Common Antigen (ECA) or a capsular polysaccharide[78]. LPS is a complex molecular species[78,79], occurring as the so-called smooth- (S-LPS) or rough-form lipopolysaccharides (R-LPS), based on the presence or absence of the polysaccharide region. Both types

consist of lipid A and, covalently linked to it, the *core* region, which is a saccharide portion, composed of up to 15 sugars[80] (Fig. 1).

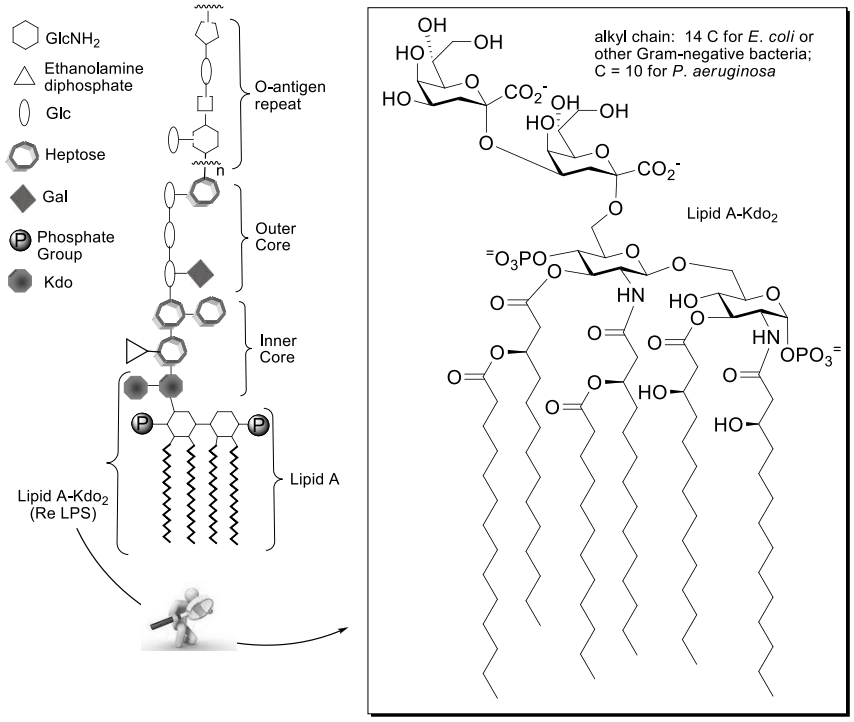


Fig. 1. Schematic view of the LPSs structure.

In the past, much emphasis has been laid on the elucidation of the structure–activity relationship (SAR). The lipid A was proven to represent the toxic principle of endotoxic active lipopolysaccharides, however, its toxicity depends not only on its structure but also on that of the core region, which is covalently linked to lipid A. LPS, also known as endotoxin, is the outermost component of the cell surface of Gram-negative bacteria. Due to the presence of LPS, the OM represents an effective permeability barrier against many toxic

compounds such as detergents, bile salts, antimicrobial peptides, antibiotics and allows Gram-negative bacteria to inhabit several hostile environments[81]. In addition, the OM contains lipoproteins, which are anchored to the OM through posttranslationally attached lipid moieties[82] and integral outer membrane proteins (OMPs), most of which span the OM *via* antiparallel β-sheet that fold into β-barrel[83] and that serve several functions, such as selective passive transport of solutes, receptors, assembly of surface structures. Within the host LPS is recognized by the CD14[84] and the Toll-like receptor (TLR)-4/MD2 complex [85] . CD14 binds and concentrates LPS, presents it to TLR-4/MD2 that triggers the biosynthesis of diverse mediators of inflammation, such as TNF-α, IL-1 and IL-6, and activates the production of co-stimulatory molecules required for adaptive immunity[79]. Lipid A is the crucial moiety in the interaction with TLR-4/MD2. Excessive response to LPS caused by the presence of circulating microbial antigens in the blood of affected patients results in severe sepsis, a rapidly progressing inflammatory disease with up to 29 % mortality[86]. Considerable efforts have been made to identify natural or synthetic molecules able to interfere with the interaction between LPS and inflammatory cells such as natural or synthetic Lipid A derivatives[87], which may operate as antagonists specifically competing for the binding to the Lipid A receptor site on TLR-4/MD2 complex. Lipid A (Fig. 1), the primary immunostimulatory *core* of LPS, is diverse in several species[88]; these variations are discriminated against during the LPS recognition

mediated by the TLR-4/MD2 complex[89]. Small differences in LPS structure have a great influence on host responses against Gram-negative bacteria[90]. For example, Lipid A of the *Escherichia coli* type acts as a potent agonist in human macrophage cells and in mouse cells. However, its precursor, lipid IVA[91], the tetra-acylated form of lipid A, acts as an antagonist in human cells but as an agonist in mouse cells[92].

On the other hand, the antibacterial design can focus not on the host-guest interaction, but on the bacteria, impairing the transport of the LPS to the OM, or its biosynthesis (Fig. 2.).

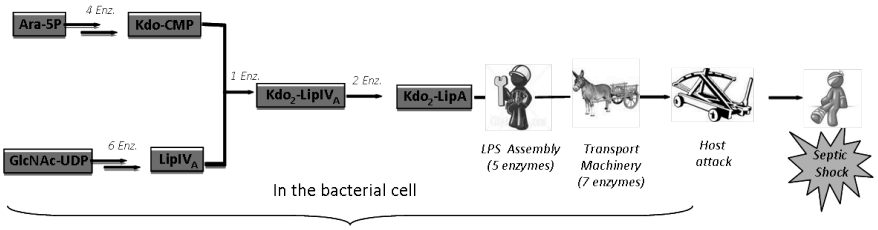
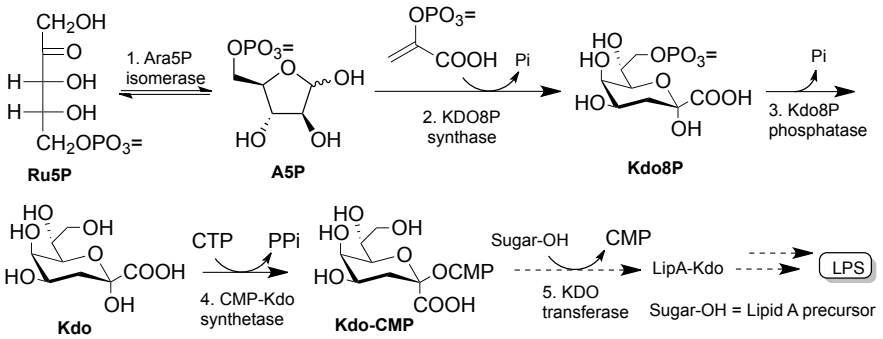


Fig. 2. Dissection of LPS biogenesis and potential targets for drug design and discovery.

Among the saccharidic residues in LPS, we highlighted 2-Keto-3-deoxy-D-manno-octulosonic acid (Kdo). Kdo is an acidic monosaccharide[93] belonging to a large family of 2-keto-3-deoxy-sugar acids, which are important constituents of complex carbohydrates, [94] playing significant roles in biological systems. Among this family, Kdo itself is an essential component of lipopolysaccharides (LPSs) in Gram-negative bacteria, and is expressed in higher plants and green algae (as a component of cell-

wall polysaccharides). Kdo was isolated for the first time in 1959[95] and in the 1960s it was recognized as a glycosidic component of the lipopolysaccharide (LPS) of Escherichia coli 0111B4.[96] Since then it has been reported in the lipopolysaccharides of all the members of the Enterobacteriaceae.[78] The significance and ubiquitous presence of this monosaccharide within LPSs promoted a through investigation into its biosynthesis and a detailed study of involved enzymes and their mechanisms. These studies are also encouraged by the observation that the breakdown of Kdo biosynthesis results in an accumulation of Lipid A precursors inside the cell, affecting bacterial viability[97]. Moreover, in enterobacterial lipopolysaccharides, the binding of the core region to lipid A occurs always *via* a Kdo residue, and as in all other lipopolysaccharides, the core region is always negatively charged (provided by phosphoryl substituents and/or sugar acids like Kdo and uronic acids) which is thought to contribute to the rigidity of the Gram-negative cell wall through intermolecular cationic crosslinks[98]. The Kdo_2-lipid A biosynthetic pathway of lipopolysaccharides[99] may be viewed as having a conserved and a variable component. The conserved (constitutive) enzymes are intracellular, present in virtually all Gram-negative bacteria, and not generally subject to regulation[79]. These observations suggest that these enzymes are excellent targets for therapeutics, possessing a broad spectrum of activity against Gram-negative bacteria[100]. In contrast, the lipid A modification enzymes, are mostly extracytoplasmic and vary from

organism to organism. Therefore, in order to inhibit the biosynthesis of LPS, we focused our attention on the biosynthesis of Kdo.



Scheme 1. Biosynthetic steps of Kdo: enzymes involved are (1) D-arabinose 5-phosphate isomerase (API, KdsD/GutQ); (2) Kdo8P synthase (Kdo8PS or KdsA); (3) Kdo8P phosphatase (KdsC); and (4) CMP-Kdo synthetase (CKS, or KdsB); (5) Kdo transferase.

Four sequential enzymatic steps (Scheme 1) are involved in the Kdo biosynthetic pathway, D-ribulose-5-phosphate (Ru5P) being the precursor:

(1) Isomerization of D-ribulose 5-phosphate (Ru5P), affording D-arabinose 5-phosphate (A5P);

(2) Condensation of phosphoenolpyruvate (PEP) and A5P to Kdo-8-phosphate (Kdo8P);

(3) Hydrolysis of the Kdo8P phosphate ester, affording Kdo;

(4) Kdo activation as a cytidine monophosphate (CMP) derivative (Kdo-CMP).

At this point biosynthetic intermediates $LipIV_A$ (in Pseudomonas aeruginosa[101] Kdo residues are transferred directly to fully acylated

lipid A) and activated CMP-Kdo converge by action of a specific membrane-bound transferase. These enzymes are all essential for E. coli survival, but when gene redundancy exists (for example when KdsD and GutQ share the same A5P isomerase activity; see later), both the isozyme genes need to be knocked out to arrest cell growth[102]. The pathway[103] is initiated by the enzyme D-arabinose 5-phosphate (A5P) isomerase (API), which converts D-ribulose 5-phosphate into A5P. The essential KdsA enzyme (Kdo 8-phosphate synthase) catalyses the second step, condensing A5P with phosphoenolpyruvate to form Kdo 8-phosphate;[104] KdsC involved in the third reaction is a phosphatase catalysing the hydrolysis of Kdo 8-phosphate to Kdo and inorganic phosphate.[105] In E. coli the kdsC gene is not essential, again indicating genetic redundancy for this function. Finally, the formation of activated sugar CMP-Kdo is catalysed by the essential KdsB enzyme.[106] Thus, transferase Waa in E. coli and in most bacteria catalyses the addition of Kdo units to lipid IVA, prior to its full acylation, while in P. aeruginosa Kdo is transferred to fully acylated lipid A.[107] Moreover, as previously mentioned, in several micro-organisms such as Haemophilus influenzae, Vibrio cholerae and Bordetella pertussis, WaaA incorporates only one Kdo residue[108]; whereas in Chlamydia trachomatis the WaaA homologue, encoded by the gseA gene, is capable of catalyzing at least three Kdo additions, consistent with the notion that Chlamydia LPS contains three Kdo residues.[109]

Over the enzymes implicated on the Kdo biosynthesis, we focused our efforts on the first enzyme, API. A5P is the first intermediate, unique to the Kdo biosynthetic pathway, a substrate that is not readily available via glycolysis. A5P is synthesised by a reversible keto–aldol isomerisation of a ketose sugar (Ru5P) to an aldose sugar (A5P), as illustrated in Scheme 1. In E. coli there are two APIs, KdsD (formerly indicated as YrbH) and GutQ,[102] which have been characterized and shown to have nearly identical biochemical properties. While all virtually sequenced genomes of Gram-negative bacteria encode KdsD, only a limited number of members of the Enterobacteriaceae also express GutQ. The biological significance of this API redundancy in E. coli is not clear yet.[110] In addition, a third paralogous gene, kpsF, has been found in pathogenic E. coli strains, such as CFT073, K1 and K5. KpsF is associated to pathogenicity islands and is implicated in polysialic capsule biosynthesis.[111,102] KdsD is the prototype isomerase, while gutQ and kpsF are paralagous genes derived from kdsD duplication and associated with other specific pathways, but are still able (at least in the case of gutQ) to substitute for kdsD. Domain analysis of KdsD indicated a core N-terminal SIS (sugar isomerase) domain (210 amino acids), followed by a pair of C-terminal CBS (cystathionine β-synthase) domains (50–60 amino acids per domain). [112] SIS domains are a usual motif in keto–aldol isomerases and also in proteins regulating the expression of genes involved in phosphosugar synthesis. Despite being common to several proteins,

the exact function of CBS domains is still unknown; a hypothesis is that they might play a regulatory role.[113] The structure of this E. coli enzyme has been recently predicted by homology modelling;[114] however, structural and catalytic details are still scarce. It is a tetrameric protein, in which each monomer presents a Rossmann fold. Predicted residues involved in the catalytic site have estimated pK_a values of 6.55 and 10.34,[115] suggesting the occurrence of a histidine or possibly a carboxylate along with a lysine or arginine at the catalytic site. Similar amino acids are known to be involved in the catalysis of other sugar isomerases,[116] indicating a common mechanism. A plausible mechanism was formulated by Rose[117] and consolidated by Walsh,[118] involving a cis-enediolate intermediate as shown in Scheme 2. A basic group on the enzyme probably removes the pro-S proton on C-1 of ribulose 5-phosphate and re-protonates the substrate at C-2 on the same (si) face of the enediol. Homology modelling studies on KdsD[114] suggest that Lys59, Glu111 and Glu152, being confined close to the presumed active site, are functional residues and are required for enzyme activity. In order to better characterise the involvement of these residues in the enzymatic catalysis, they were mutated to alanine: the Lys59Ala and His193Ala mutants showed a drastic loss of activity, indicating these two residues as essential players in catalysis. Moreover, NMR epitope studies performed on natural substrates and analogues (Fig. 3) highlighted key features needed for enzyme recognition and keto–aldol interconversion.[119]

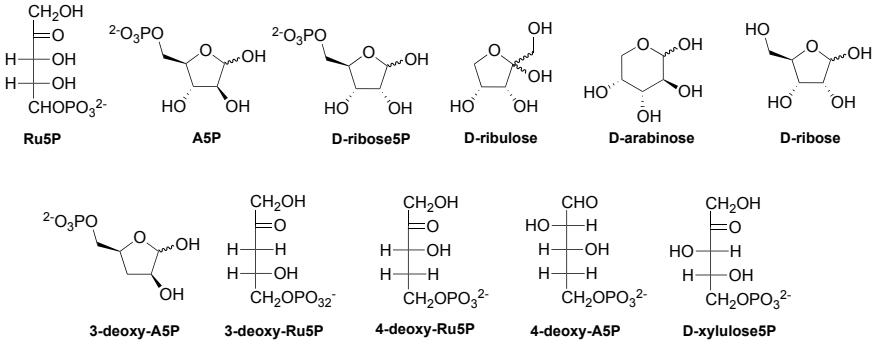
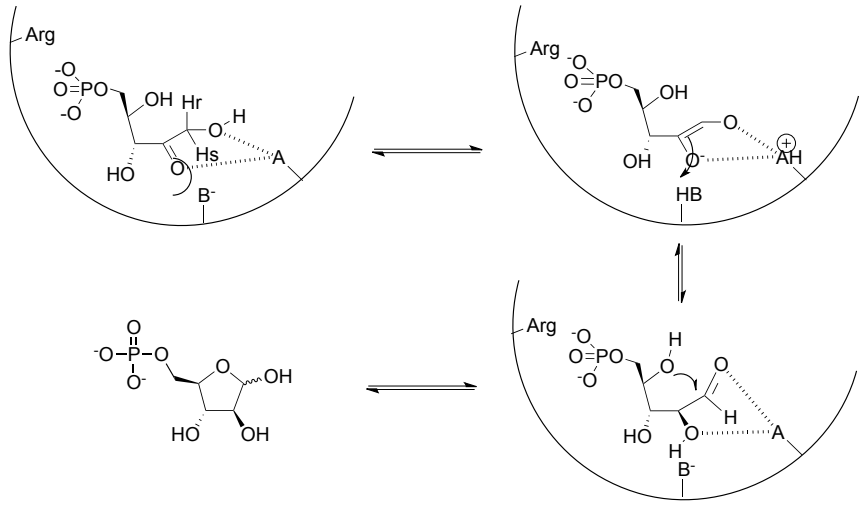


Fig. 3. Natural substrates and analogues used for NMR studies.

The acidic phosphate group at position 5 is fundamental for enzyme recognition; the OH in position 3 is important for the recognition, and in particular its stereochemistry is fundamental. On the other hand, substituents at positions 2 and 4 do not seem to be essential for the interaction with the enzyme. A hydroxyl group at position 4 (which is involved in the hemi-acetal ring formation) is not needed for recognition or for isomerisation. Rates indicate that keto–aldol isomerisation is faster on substrates such as 4-deoxy-A5P, thus suggesting that substrate in the open-chain form speeds up the reaction, provided that the remaining functional groups are in place (i.e. phosphate and 3-OH). In addition, the effect of Zn^{2+} was investigated. It has been reported that late-transition metals (Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, Hg^{2+}), in particular those with the d_{10} electron configuration, inhibit KdsD.[112] According to NMR data, Zn^{2+} inhibits substrate binding in a reversible manner. [119]



Scheme 2. Postulated enzyme mechanism for A5P isomerase.

Starting from these data I designed A5P derivatives that potentially can inhibit API, or however can interfere with the Kdo biosynthesis. In order to reach this goal, arabinose has been used as the scaffold, instead of ribulose, because arabinose is not metabolised by humans.

Over the years, several analogues of Kdo biosynthetic intermediates have been designed[77]: some of them showed potent inhibitory activity *in vitro,* however, none of them turned out to be promising *in vivo* due to their inability to permeate the bacterial membranes.

It is reasonable to speculate that also A5P isomerase inhibition might have similar cellular effects observed for Kdo synthase inhibition, thus making this enzyme a valuable drug target. Based on the hypothesised enzymatic mechanism (the mechanistic details still have to be elucidated), a number of inhibitors were synthesised three

decades ago[77]: different classes of arabinose 5-phospahte mimics have been suggested as inhibitors, such as A5P analogues, aldonic acid phosphates, alditol phosphates, and a series of non-phosphorylated arabinose derivatives and ribonolactones, with scarse results. Recently Woodard and co-workers[120] chose to explore hydroxamic acids, based on the fact that they have been already used to mimic the enediol intermediate in similar systems. So they tested the most potent inhibitors of *T. brucei* 6PGDH; compounds **1**-**3** (Fig 4) shown activity on *F. tularensis* KdsD (IC_{50} 35, 7 and 10 μM), but when they were tested in vivo in *E. coli*, only **2** and **3** were active in the millimolar range (IC_{50} 52 and 18 mM). However, the limited selectivity and cytotoxicity data available, suggest that work will be required to reach a solid lead compound.

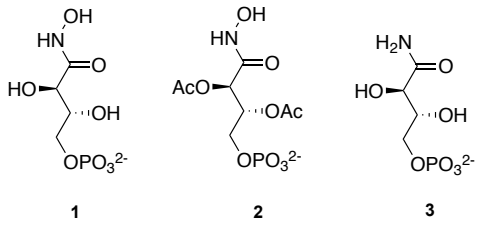


Fig. 4. The most potent API inhibithors to date.

3.2. API: Results and discussion

In order to better characterise A5P isomerase and influence the LPS biosynthesis in gram-negative bacteria, I designed and synthesized three sets of glycomimetics, always using arabinose as the scaffold:

1) Deoxy-A5P analogues. Mimetics of arabinose-5-phosphate, lacking a hydroxyl group in position four or in position three.
2) A5P Cyclic analogues. A5P derivatives blocked into their cyclic forms as methyl arabinosides.
3) 5-modified A5P analogues. A5P mimetics, modified only in the acidic group in position 5.

1) Deoxy-A5P analogues. Previous *in vitro* studies on isolated API[119] showed that both the compounds are recognized and processed by the enzyme. However, API has a lower specific activity for 3-deoxy-A5P ((3.295 ± 0.165) x 10^{-5}) compared to A5P ((2.306 ± 0.467) x 10^{-4}), while 4-deoxy-A5P has a higher affinity for the enzyme ((4.187 ± 0.491) x 10^{-4}) compared to the natural substrate. Both these A5P analogues had never been tested *in vivo*, so I decided to synthesize them in order to test their biological activity because even if we already know that they cannot be inhibitors of API, these compounds have interesting features in light of the Kdo biosynthetic pathway.

3-deoxy-A5P could be the inhibitor of the enzyme subsequent to API (Kdo8P), but if it is recognized and processed from Kdo8P anyway because of the lacking 3-OH, it will afford in a linear form of Kdo8P (Fig. 5a) that presumably cannot go further in the LPS biosynthesis.
On the other hand, Kohen.[121] observed that 4-deoxy-A5P *in vitro* is substrate of Kdo8P, giving a modified Kdo8P, deoxy in its position 7 (Fig. 5b); this Kdo8P analogue could potentially cause the inhibition of the following enzyme. Anyway if it is processed, it could lead theoretically to the formation of modified LPS, interesting for our studies.

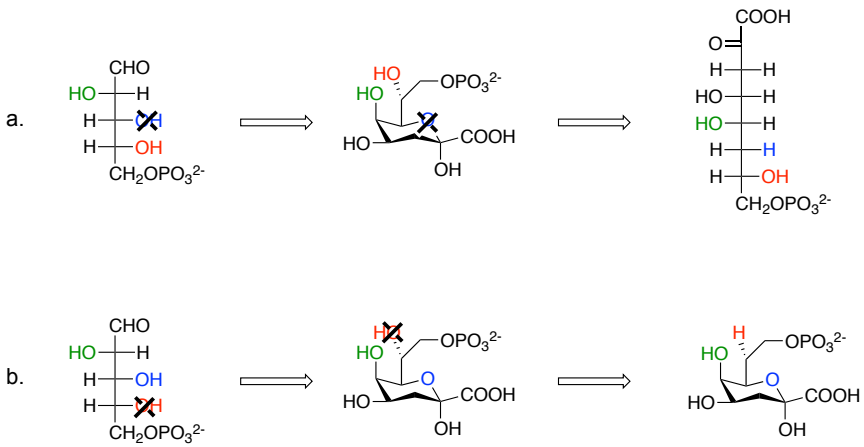
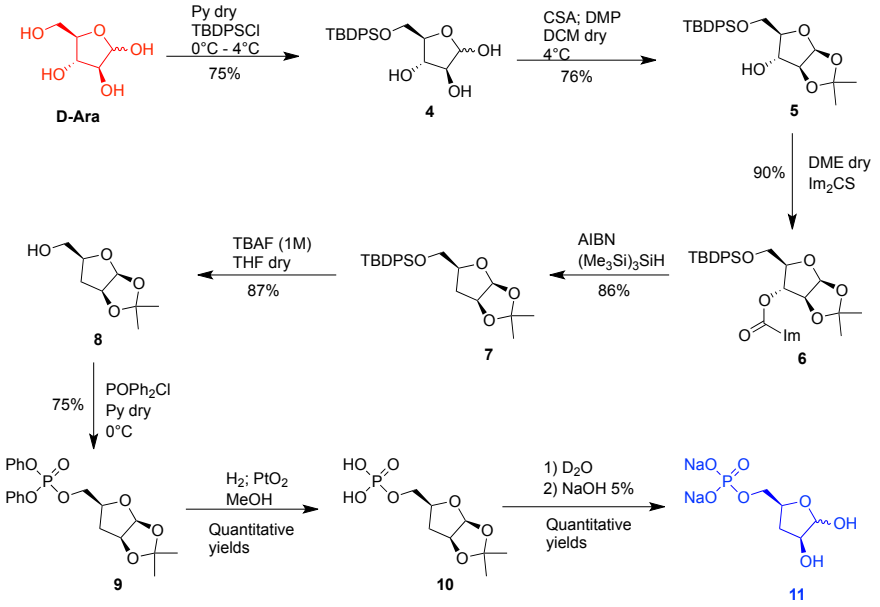


Fig. 5. Potential modified Kdo obtainable starting from 3-deoxy-D-A5P (a) and 4-deoxy-D-A5P (b).

The synthesis of 3-deoxy-D-A5P[122] was performed starting from D-arabinose (Scheme 3), which was treated first with *t*-butyldiphenylsilyl chloride (TBDPSCl) in order to selectively protect the primary hydroxyl group affording compound **4**, and then with

acetone dimethylacetal in dichloromethane, in the presence of camphorsulphonic acid (CSA), to protect the hydroxyl groups in positions 1 and 2. These two hydroxyl groups are the only cis-related, in the β-anomer, that can generate the isopropylidene cycle. The obtained *5-O*-t-butyldiphenylsilyl-1,2-*O*-isopropylidene-β-D-arabinofuranose (**5**) has the only hydroxyl group at C-3 deprotected, allowing its reduction to hydrogen, exploiting the radical displacement of a thiocarbamate. For this purpose compound **5** was treated with thiocarbonyldiimidazole in dry dichloroethane at 80°C to afford the carbamate **6**, which was directly treated with $(Me_3Si)_3SiH$ and AIBN in toluene under reflux, affording *5-O*-t-butyldiphenylsilyl-3-deoxy-1,2-*O*-isopropylidene-β-D-*threo*pentofuranose **7** in *77%* yield over two steps. The deoxygenated compound **7** was then deprotected at the primary hydroxyl group by treatment with tetrabutylammonium fluoride (TBAF) in THF affording 3-deoxy-1,2-Oisopropylidene-β-D-*threo*pentofuranose **8** in *87%* yield. The primary hydroxyl group of **8** was then converted into a diphenylphosphate **9** by treatment with diphenyl chlorophosphate in dry pyridine (*75%* yield), and the diphenylphosphate was quantitatively deprotected by treatment with H_2/PtO_2 in degassed methanol, affording 3-deoxy-1,2-*O*-isopropylidene-β-D-*threo*pentofuranose 5-phosphate **10**. Finally, hydrolysis of the isopropylidene, the last protecting group, afforded 3-deoxy-D-*threo*pentofuranose 5-phosphate **11** in quantitative yield.

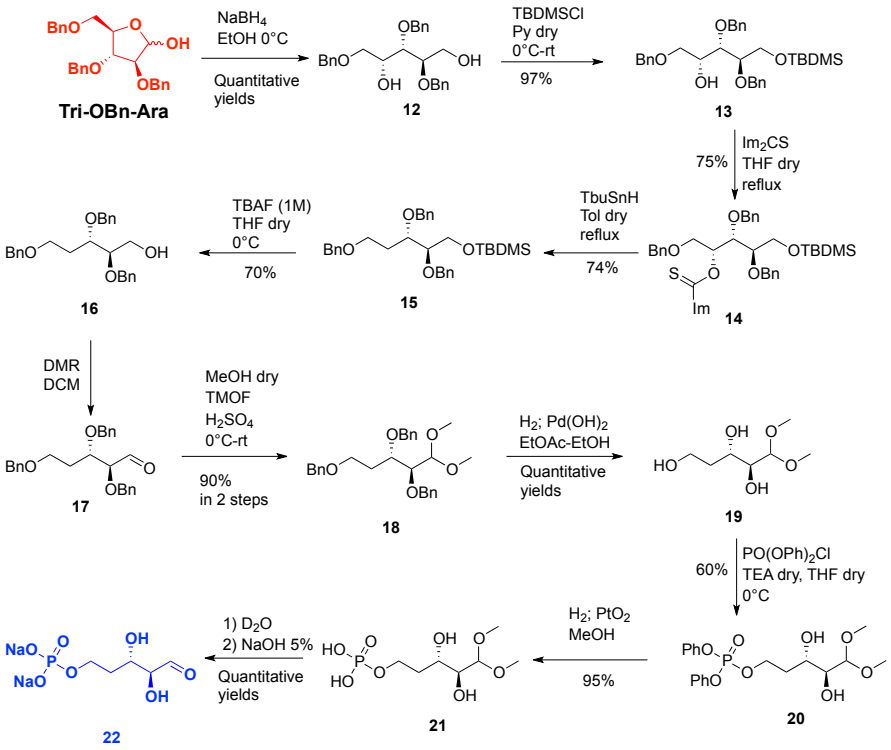


Scheme 3. Synthesis of 3-deoxy-A5P.

4-deoxy-D-A5P[121], the open-chain analogue of D-arabinose-5-phosphate, was prepared from commercially available 2,3,5-tri-O-benzyl-D-arabinofuranose, as it is shown in Scheme 4.

Reduction with $NaBH_4$ (12) and a selective protection of primary alcohol in position 1 with *t*-butyldimethylsilyl ether afforded the protected derivative 13. At this point the only position available for modification is position 4, which has been deoxygenated by conversion of the hydroxyl group to the O-(imidazoylthiocarbonyl) derivative (14) followed by treatment with tri-n- butylstannane, (15) in an overall yield of 52% for two steps. Desilylation with TBAF in dry THF then provided the corresponding primary alcohol (16), whose oxidation with the Dess-Martin periodinane in dichloromethane to give the aldehyde 17. The aldehyde was

converted to the dimethylacetal **18** by treatment with trimethyl orthoformate, followed by benzyl groups removal (in EtOH, over $Pd(OH)_2$, **19**) and selective phosphorylation of the primary hydroxyl group with diphenyl-phosphochloride to give the protected phosphate monoester **20**. Catalytic hydrogenolysis of **20** over PtO_2, in MeOH yielded the corresponding phosphoric acid **21**. This aldehyde is quantitatively deprotected in a D_2O solution into a NMR tube and finally the pH is neutralized by the addition of 5% solution of NaOH, giving the corresponding disodium salt **22**.



Scheme 4. Synthesis of 4-deoxy-A5P.

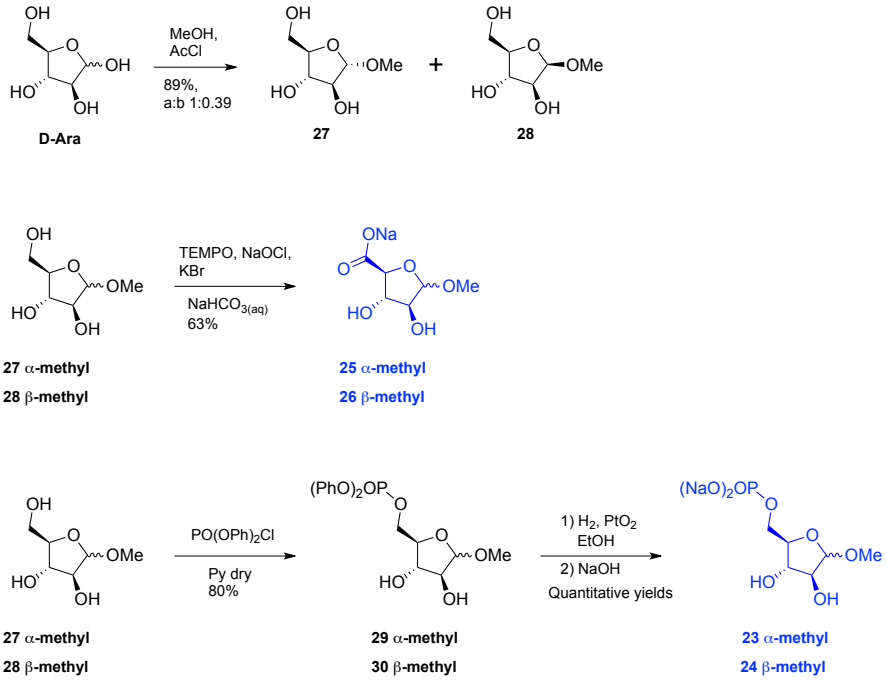
At this point the two synthesised deoxy analogues were tested *in vivo* on three different strains of *E. coli*: the wild type strain BW25113, the permeable strain AS19 [123] and on the BB8 mutant[123], but no biological activity was observed.

2) A5P Cyclic analogues. A5P isomerase catalyses in a reversible way the isomerization between Ru5P and A5P; the ketose ribulose is obviously linear, while arabinose is in equilibrium between the linear form and the cyclic emiacetalic form. Thus, we know that API is able to recognize the substrate in the linear form, and the STD studies on 4-deoxy-A5P (linear) suggest that API has a higher affinity for linear substrates. However, we do not have any information on the ability of the enzyme to bind cyclic substrate. In order to answer this question we decided to block the emiacetalic form as methyl arabinosides, synthesising derivatives **23** and **24**, maintaining the phosphate group in position 5 required for binding. Moreover, we synthesised the corresponding analogues **25** and **26**, in order to check if the non-hydrolysable carboxylate moiety could be a good phosphate mimetic.

To synthesise the targeted arabinosides (**23-26**), we first performed the synthesis of the methyl arabinofuranosides **27-28**. We subjected D-arabinose to a solution of AcCl in MeOH and stirred the reaction for 18 h at room temperature[124] (Scheme 5). As already observed by Dangerfield and co-workers, these conditions led to the formation of the desired methyl furanosides **27-28** in 89% yield,

instead of the undesired thermodynamically more stable methyl pyranoside. The α and β anomers (α/β 1:0.39) were separated and used as intermediates for the synthesis of the phosphate and carboxylate analogues. To achieve the synthesis of the phosphate moiety, we first phosphorilated both the α and β methyl D-arabinofuranoside (compounds **27** and **28**) using diphenyl chloro-phosphate and pyridine as solvent, obtaining compounds **29** and **30** in good yields (around 80%). Next we removed the phosphate's protective groups by hydrogenolysis reaction, using Platinum oxide as catalyser; subsequent pH neutralization with sodium hydroxide gave the desired cyclic phosphate analogues **23** and **24** as sodium salt in quantitative yields. The neutralization step is fundamental, because freeze-drying the aqueous solution at its natural acidic pH brings to a mixture of compounds, mainly due to the hydrolysis of the methyl group.

The cyclic carboxylate analogues **25** and **26** were synthesised oxidising methyl D-arabinofuranoside (**27**-**28**) using sodium hypochlorite and TEMPO as oxidant at catalytic amount (around 63% yields).



Scheme 5. Synthesis of A5P cyclic analogues.

The phosphate and the carboxylate cyclic analogues were studied by STD-NMR experiments on the isolated API.

The binding of compounds **23**-**24** was tested by STD-NMR working on a mixture containing both the anomers at a 1:1 molar ratio; in this way, we were able to directly compare their affinity for API on the basis of the relative STD intensities. The ^{1}H NMR spectrum of a mixture containing the α and β anomers and the enzyme dissolved in PBS, pH 7.0, 37°C, is depicted in Fig. 6. This spectrum was acquired about 3 h after incubation of **23** and **24** with the protein and, as expected as a consequence of position 1 glycosilation, clearly showed that they are not API substrates, as

their isomerization to the corresponding ketoses did not occur. Fig. 6 reports the STD spectra recorded on the same mixture and indicated that both the glycosides are API ligands. In addition, β anomer has lower affinity for the protein, as its STD signals are less intense than those of α anomer.

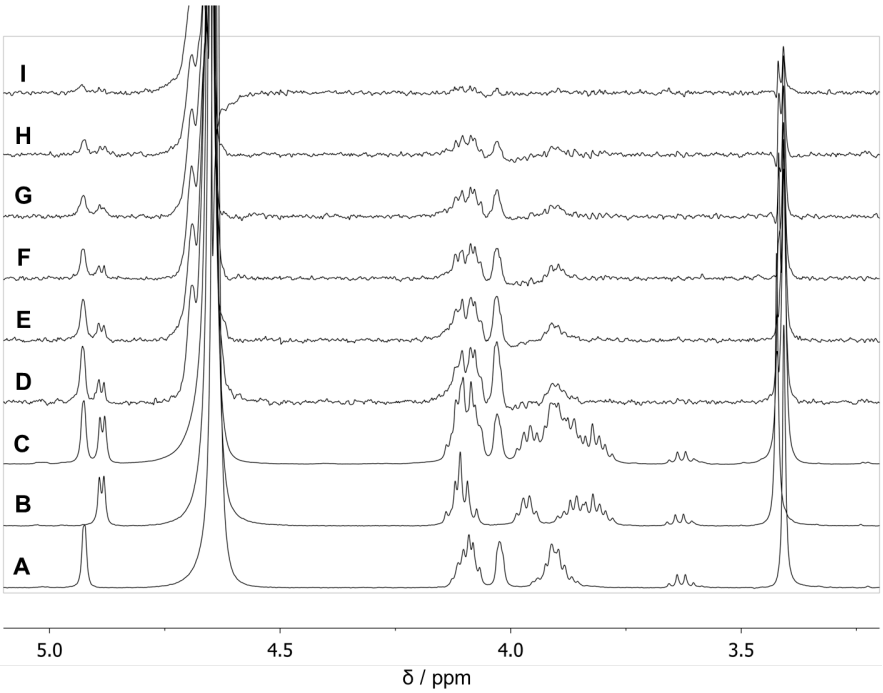


Fig. 6. A) ^{1}H NMR spectrum of a solution of compound 23 (3 mM); B) ^{1}H NMR spectrum of a solution of compound 24 (3 mM); C) ^{1}H NMR spectrum of a mixture containing of compound 23 (3 mM), compound 24 (3 mM) and API (30 uM); D-I) STD-NMR spectra of the a mixture containing of compound 23 (3 mM), compound 24 (3 mM) and API (30 uM) acquired with different saturation times (D, 3.0 s; E, 2.0 s; F, 1.5 s;

G, 1.0 s; H, 0.6 s; I, 0.3 s) and a saturation frequency of 7.9 ppm. All samples were dissolved in PBS, pH 7.0, 37°C.

To verify if the recognition of compounds **23** and **24** occurred at API binding site, a competitive STD with natural substrates A5P and Ru5P was performed. In particular, the competitive STD experiment was performed on a mixture containing API (30 uM), the natural substrates (total concentration 5 mM) and compounds **23** and **24** (3 mM each). Signals of **23** and **24** are almost absent in the competitive STD spectrum (Fig. 6F), while natural substrate resonances appear clearly, indicating that the glycosides are displaced by A5P and Ru5. This evidence demonstrated that molecules **23** and **24** bind the enzyme in the active site, also with a lower affinity if compared with natural substrates.

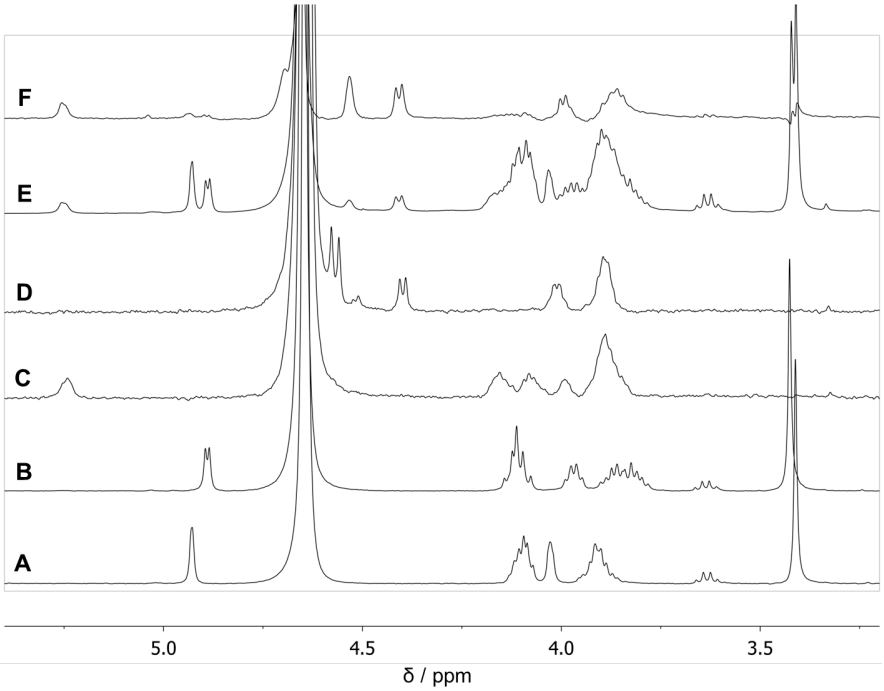


Fig. 7. A) ^{1}H NMR spectrum of a solution of compound **23** (3 mM); B) ^{1}H NMR spectrum of a solution of compound **24** (3 mM); C) ^{1}H NMR spectrum of a solution of A5P (3 mM); D) ^{1}H NMR spectrum of a solution of Ru5P (3 mM); E) ^{1}H NMR spectrum of a mixture containing of compound **23** (3 mM), compound **24** (3 mM), A5P and Ru5P at equilibrium (total concentration 5 mM) and API (30 uM); F) STD-NMR spectra acquired on the same mixture of spectrum E with a saturation time of 2.0 s and a saturation frequency of 7.9 ppm. All samples were dissolved in PBS, pH 7.0, 37°C.

These data suggest that probably API is able to recognize the natural substrates also in hemiacetalic form; nevertheless, the cyclic form of the ligand and/or the introduction of a substituent in anomeric position dramatically reduce the affinity.

Compound **25** and **26** were tested as API ligands according to the same experimental approach. STD data indicated that the enzyme

does not recognize these compounds. Thus, the carboxylate group does not properly mimic the phosphate group present in the natural substrates, at least regarding its activity on API.

To better investigate the properties of compounds **23**-**26**, we tested their ability to inhibit the growth of *P. aeruginosa* in liquid culture. As Pa-KdsD shares a high level of sequence similarity with its *E. coli* counterpart, we tested an *E. coli* mutant strain (AS19) that is more permeable to many antibiotics[123], in order to address the extent to which the Gram-negative bacteria cell barrier limits the action of the compounds. In addition, we also assayed the effect of the compounds on growth of wild type strain of *E. coli* (BW25113) in the presence or absence of glucose-6-phosphate (G6P), since G6P induces the hexose phosphate transport system (*uhp*) in *E. coli* and A5P is a high-affinity, though non-inducing, substrate of this system.[102]

As shown in Table 1, none of the compounds appeared to exert inhibitory growth effect neither on *P. aeruginosa* nor on wild type *E. coli* strain at the concentrations tested. Epitope mapping studies revealed that compounds **23** and **24** interact with Pa-KdsD, and that this interaction occurs at the active site of the enzyme; however, their binding is weaker than the natural substrates. These results are in line with the slight inhibitory activity observed *in vivo* on bacterial cultures of the permeable mutant AS19 where the minimal inhibitory concentration (MIC) of each anomers was 250 μM. Surprisingly, also compounds **25** and **26**, that were not recognized by the

enzyme in STD-NMR experiments, led to growth arrest of AS19 cultures with a MIC of 250 μM. This suggests that also compounds **25** and **26** function as weak inhibitors, possibly by interacting with the Kdo 8P synthase, the enzyme acting downstream KdsD in the Kdo biosynthetic pathway[112].

Nevertheless, an intact outer membrane normally prevents the entry of these four compounds into the cell, as no inhibitory activity was observed against wild type *E. coli* and *P.aeruginosa*. This suggests that rounds of optimization are still required to improve cellular uptake of these compounds.

	MIC, mM			
	Compound:			
Strain	23	24	25	26
P. aeruginosa PA01	>2	>2	>2	>2
E. coli BW25113	>2	>2	>2	>2
E. coli BW25113 + G6P	>2	>2	>2	>2
E. coli AS19	0,25	0,25	0,25	0,25

Table 1. MIC determination of compounds **23-26**

3) 5-modified A5P analogues. Previous NMR studies on API[119] showed that the phosphate group is fundamental for the binding with the enzyme, and in particular the salt bridge that occurs between the negatively charged phosphate and presumably the

positive Arg residue in the catalytic pocket, is fundamental [125]. Due to the poor stability of the phosphate group to phosphatase[126], the finding of a good and chemically and enzymatic stable phosphate mimetic is a key point for the design of potential inhibitors of API.
Thus, I designed and synthesised a small library of A5P analogues, changing only the acidic group in position 5 (Fig. 8).

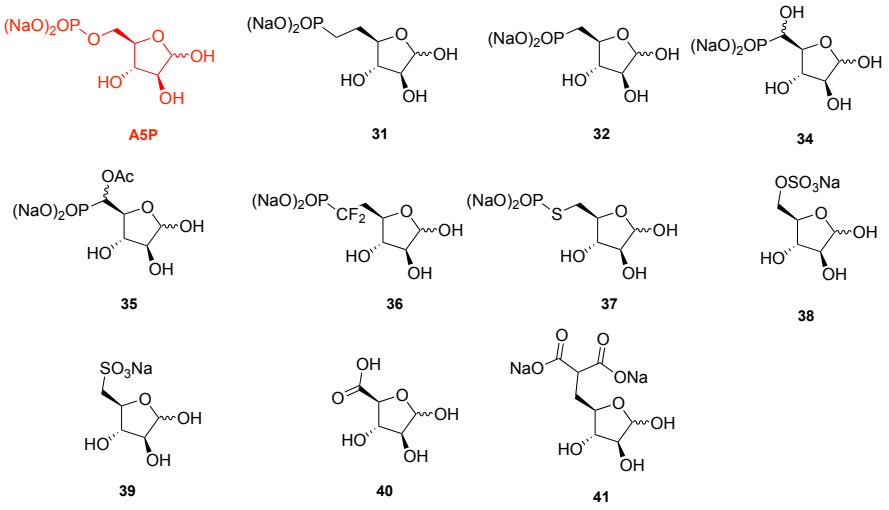


Fig. 8. Set of A5P analogues with different acidic group at position 5.

This set of compounds also has an interesting feature because it can be potentially biologically active in different ways: these A5P analogues could be inhibitors of both API or Kdo8P, but if they are recognized and processed because the acidic group is a good mimetic of the natural phosphate, they cannot be hydrolysed by the Kdo8P phosphatase and they cannot proceed further into the LPS biosynthesis (Fig.9).

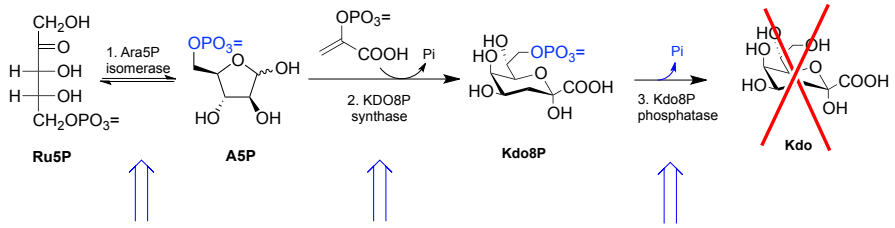


Fig. 9. Three possible targets for A5P analogues with modified acidic group in position 5. The natural phosphate group is highlighted in blue.

The natural phosphate group has substituted in compound **31** with the isosteric methylphosphonate. Several aspects of change must be considered on the performance of this fundamental structural substitution, aside from rendering the phosphorus not liable to hydrolysis by normal routes. First, there would be expected to result some significant decrease in acidity of the phosphorus-containing acid function upon the introduction of an electron-donating alkyl group. This could result in the existence of a different state of dissociation for the analogue compared to the natural compound at a particular (physiological) acidity associated with a biological system[127]. It should be noted that while the term "isosteric" strictly refers to compounds of identical size and shape[127], and since **31** compared to A5P do not meet these conditions most rigorously, the bond angles and lengths involved are similar enough that the term reasonably may be applied. Unger [128] in 1978 synthesized compound **31** and tested it only *in vitro* on isolated Kdo8P synthase (the enzyme subsequent to API); he observed that the A5P analogue is substrate of the enzyme, and the product of the enzymatic

reaction inhibits the activity of Kdo8P synthase. However, no *in vivo* test and *in vitro* assay on API have been ever performed. Thus, I synthesized **31** performing a new synthetic scheme that allows to obtain the desired analogue in less synthetic steps (9 instead of 13) and with a higher total yield (28% instead of 18%) compared to what is reported in the literature.

In order to explore the importance of the distance between the negative charges and the furanose ring, I tried to synthesise the non-isosteric phosphonate **32** (Fig. 8) through, a Michaelis-Arbuzov reaction on the iodo-derivatived **33** (Scheme 11), but without success. Thus, to study the effect of a non-isosteric phosphonate I successfully synthesised the α-hydroxy-phosphonate **34**. Moreover, the presence of an additional hydroxyl group in position 5, allows the compound to be present in equilibrium between its furanosidic and pyranosidic form. This can also give us additional information about the ability of API to recognize the pyranose ring.

To avoid the formation of the pyranosidic form, and to try to isolate the effect of the furanosidic form, I synthesised compound **35** that has the hydroxyl group in position 5 acetylated; moreover, the presence of the acetyl group in α to the phosphonate, would presumably slightly increase its acidity.

Difluoromethylenephosphonates are postulated[126] as stable biological models of natural phosphates. This is associated with the fact that the CF_2 group is an isoelectronic and isosteric analogue of the oxygen atom, wherein the fluorine atoms are considered as

oxygen lone pairs. In the last 15 years a wide range of biologically active compounds containing the difluoro-methylene groups, which are derivatives of nucleosides, amino acids, sugars and heterocycles, were synthesized. These compounds are of interest as potent inhibitors of various enzymes and as probes for the elucidation of the mechanisms of biological processes with the participation of phosphates. The spatial parameters are similar to those of the phosphate[129], moreover, the presence of two fluorine atoms will increase the polarity of the group, increasing the acidity of the phosphonate (second pKa 5.4, Phosphate 6.4), while the methylenphosphonate has a second pKa lower than the natural phosphate one. Although fluorine is the most abundant halogen element and ranks number 13 among all the elements in the earth's crust, the naturally occurring organofluorine compounds (compounds containing one or more C–F bonds) are rare. Among nearly 3,200 known naturally occurring organohalogen compounds, only very few (around 13) belong to organofluorine compounds (all of them are monofluorinated!)[130]. This makes the presence of two fluorine atoms interesting also for using this analogue as molecular probe. Thus, I synthesised the corresponding 5-difluoromethylenephosphonate-arabinofuranoside **36**, by substitution of a triflate intermediate, with nucleophile litium-difluoromethylenephosphonate, as further explained in detail.

The fifth phosphorous-based mimetic of A5P that I synthesized is the thiophosphate analogue **37**. Among the phosphate esters,

phosphorothioate derivatives are of interest as effective pesticides; in recent years a number of phosphorothioates have been introduced as potential chemotherapeutic agent [131]. The phosphorothioate analogs of phosphate esters have been useful tools in the study of the mechanisms of several enzymes catalysing phosphate ester hydrolysis[132]. The differences in enzymatic reactivity observed can be attributed to alterations in the physical properties and structural features of the ester. For example, the reduction in polarization of the phosphorus by the less electronegative sulphur atom is an important peculiarity. In addition, the interatomic distance is also considerably larger than oxygen because of the higher Van der Waals radius of sulphur. These structural differences may have important consequences in the formation of the enzyme-substrate complexes, as well as effecting perturbations of certain steps in the enzymatic pathway of hydrolysis.

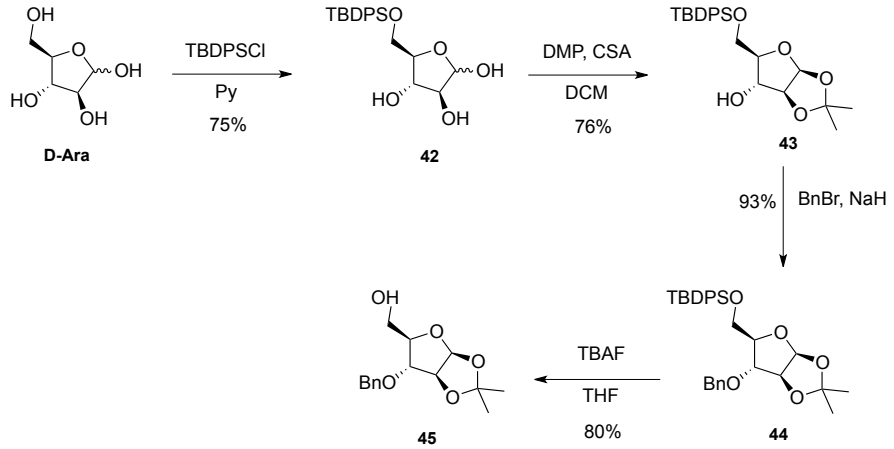
A part the phosphorous-based analogues, I also explored sulphur-based analogues, such as the sulphate **38** derivative. D-Ribono-1,4-lactone-5-(sodium sulphate) showed a very low inhibitory activity on API [133], however the sulphate group has never been used as phosphate mimetic on arabinose scaffolds.

The non-isosteric sulphonate **39** has been synthesized as well, in order to compare its activity to the sulphate's activity.

Carboxylate analogues have been explored, such as the non-isosteric carboxylate **40** that, as previously reported, in methylarabinofuranosides **25-26**, showed bacteriostatic activity.

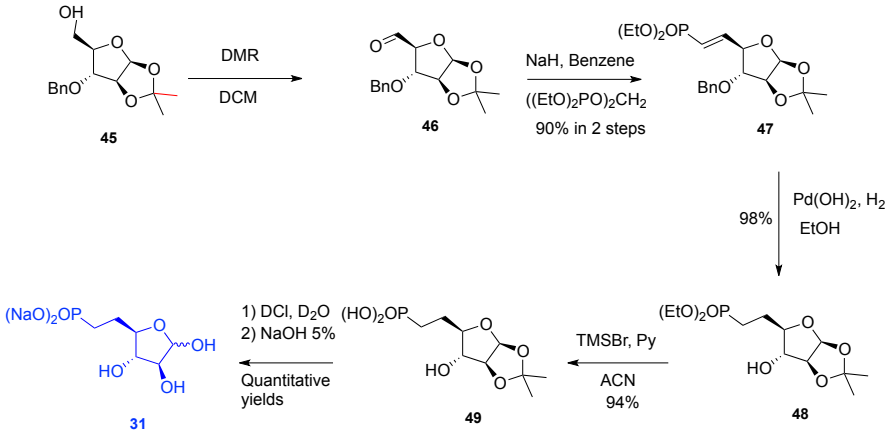
To check the effect of a double charge on a carboxylate moiety I synthesized the malonate analogue **41**. It is not proper to define it as phosphate isosteric group; however, they have the same number of atoms (3 atoms) between the negative charge and the arabinosidic C-5; in addition, this compound can display two negative charges as the phosphate group. Examples where the malonate is a good mimetic of the phosphate group are already present in the literature; for example, it is reported[134] that malonate analogue of mannose-6-phosphate bind the CI-M6P receptor with the same affinity of the natural substrate.

The synthesis of 5-modified A5P analogues has been performed starting from D-Arabinose (Scheme 6), which was first treated with *t*-butyldiphenylsilyl chloride (TBDPSCl) in order to selectively protect the primary hydroxyl group affording compound **42**, and then with acetone dimethylacetal in dichloromethane, in the presence of camphorsulphonic acid (CSA), to protect the hydroxyl groups in positions 1 and 2 (**43**). Benzylation of the hydroxyl group in position 3 (**44**) and subsequent selective cleavage of the sylilether with TBAF, gave compound **45**, a key intermediate, because the only position available for modification is position 5; therefore, starting from this compound I synthesized most of the 5-modified A5P analogues.



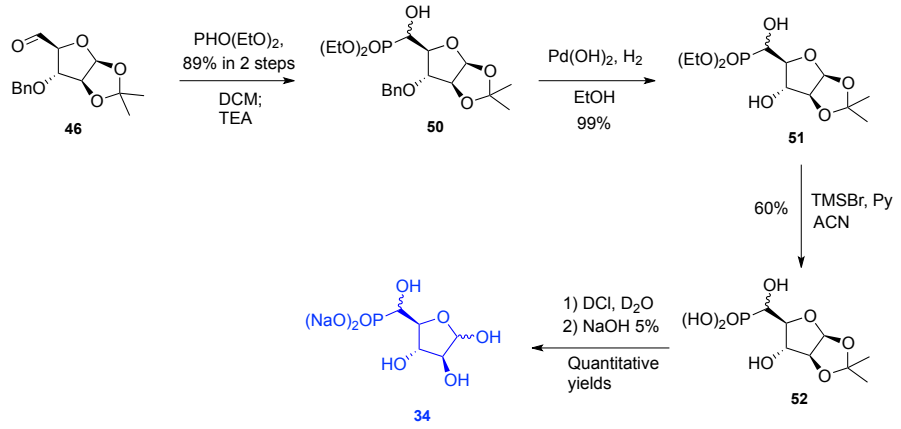
Scheme 6. Synthesis of key intermediated for the synthesis of 5-modified A5P analogues.

After oxidation of the alcoholic function with Dess-Martin reagent in DCM, I obtained the corresponding aldheide **46**; subsequent treatment with methyl-bis-diethylphosphonate[135] and sodium hydride in dry benzene gave the unsaturated phosphonate **47**. Hydrogenolisys reaction catalysed by platinum oxide in EtOH cleaved benzyl group and reduced the double bound **48**. The ethyl group were deprotected by TMSBr in pyridine and acetonitrile. The hydrolysis of the isopropylidene, the last protecting group, was performed in a D_2O solution of **49**; final pH neutralization afforded the methylenphosphonate analogue **31** in quantitative yield.



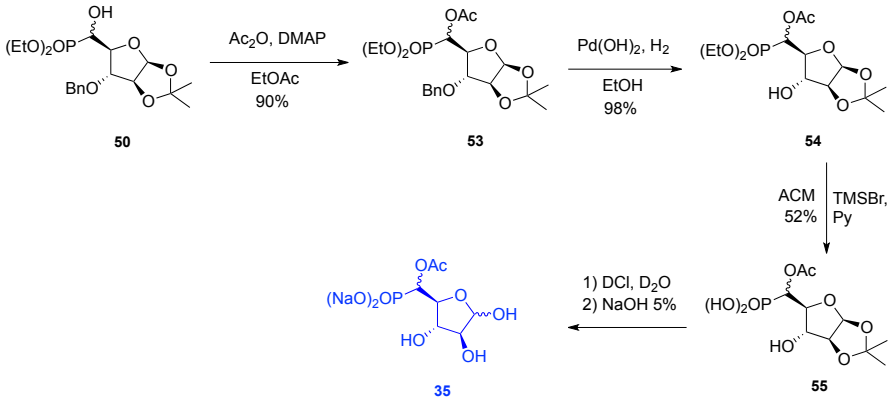
Scheme 7. Synthesis of D-Ara-5-methylphosphonate.

Reaction of the freshly prepared aldehyde **46** with diethylphosphite [136] in DCM and TEA gave the α-hydroxy-phosphonate **50** (Scheme 8). Subsequent cleavage of the benzyl group in EtOH and $Pd(OH)_2$ (**51**) and deprotection of the phosponhate ethyl groups with TMSBr in acetonitrile and pyridine gave compound **52**. Hydrolysis of the dimethylacetal and final pH neutralization gave the final 5-α–hydroxy-phosphonate **34** in quantitative yield.



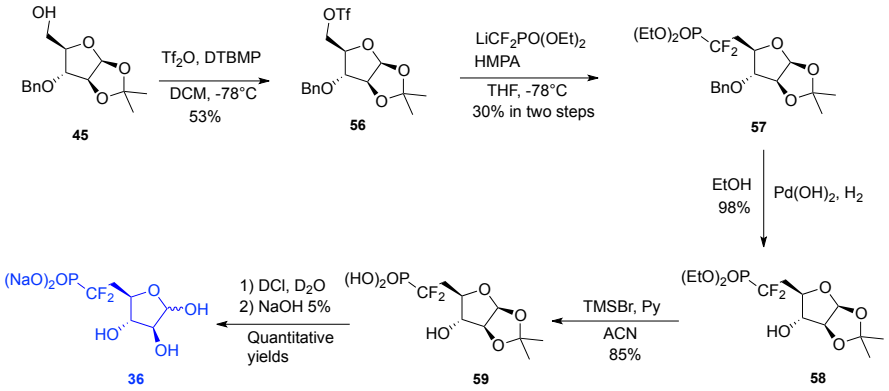
Scheme 8. Synthesis of α-hydroxy-phosphonate.

Acetylation of the free hydroxyl group of compound 50 gave the corresponding α–acetyl-phosphonate 53 (Scheme 9). Cleavage of the benzyl group by hydrogenololysis (54) and deprotection of the ethyl groups with TMSBr in acetonitrile and pyridine afforded compound 55. Hydrolysis of the dimethylacetal and final pH neutralization gave the final unstable α-O-acetyl-phosphonate 35.



Scheme 9. Synthesis of α-OAcetyl-5'-phosphonate.

Compound 45 was treated with triflyc anhydride (Tf_2O) and diterbutyhl-methyl-pyridine (DTBMP) because pyridine is known to replace the O-triflate [137], obtaining the triflate intermediate 56 (Scheme 10). This compound is not stable, so it is immediately reacted with diethyldifluoromethylphosphonate [138] using lithiodiisopropylamine as base, in presence of HMPA in dry THF: displacement of the triflate with diethyldifluoromethylphosphonate gave compound 57. Subsequent cleavage of the benzyl group in EtOH and $Pd(OH)_2$ (58) and the ethyl groups with TMSBr in ACN and Py afforded compound 59. Deprotection of the isopropylidene was performed in a D_2O solution and after final pH neutralization afforded methylendifluorophosphonate 36 in quantitative yield.

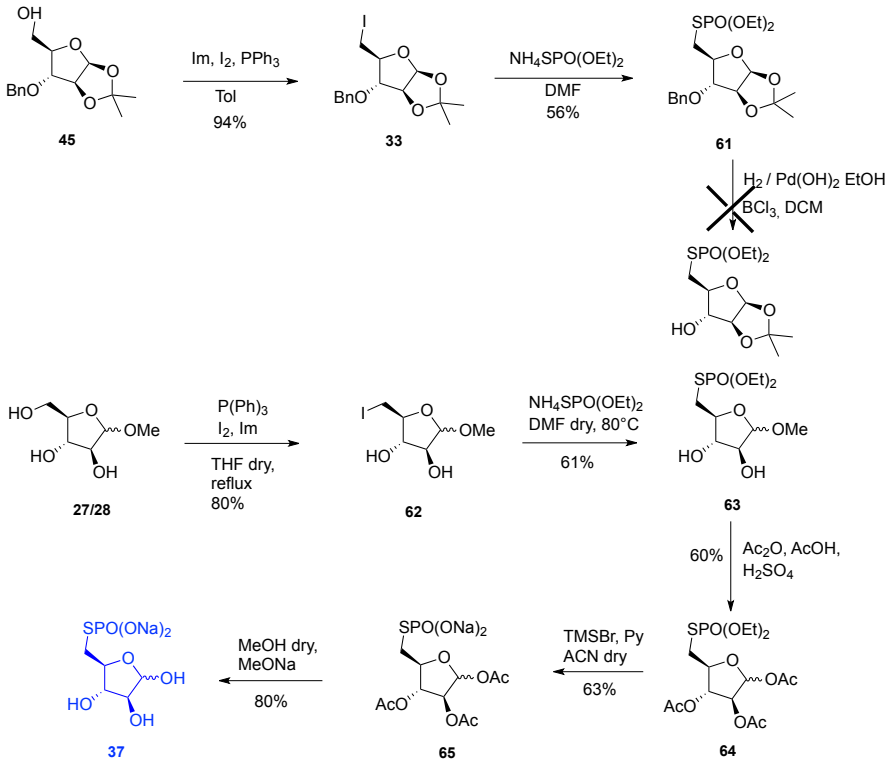


Scheme 10. Synthesis of methylendifluorophosphonate analogue.

The first attempt to synthesize the thiophoshate analogue (Scheme 11), was done synthesizing the 5-iodo-3-benzyl-1,2-O-isopropylidene-β-D-*threo*pentofuranose by Mitzunobu reaction on compound **45** with triphenylphsphine, imidazole and iodine in dry toluene, affording the iodo derivtived **33**. Substitution of the iodine in DMF with ammonium diethylthiophosphate, that I previously synthesised reacting sulfur with diethylphosphite in the presence of ammonium hydrogen carbonate[131], gave the corresponding phosphorothioate **61**. Unfortunately, the cleavage of the benzyl group in position 3 was unsuccessful by hydrogenolysis reaction catalysed by $Pd(OH)_2$ and also by reaction with BCl_3 in DCM. Therefore, I changed synthetic scheme, in order to avoid the presence of the benzyl ether protective group.

Starting from D-arabinose, I synthesized the corresponding methyl arabinofuranoside as an α/β mixture (**27**/**28**) subjecting D-arabinose to a solution of AcCl in MeOH[139] as previously reported.

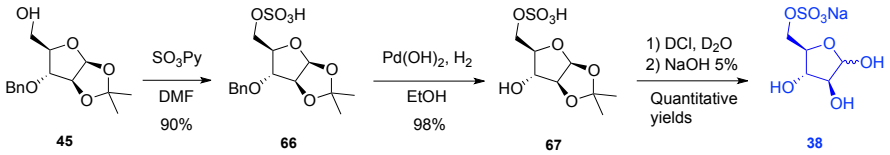
A selective Mitzunobu reaction on the primary alcohol with triphenylphsphine, imidazole and iodine in dry toluene, afforded the 5-iodo-methyl arabinofuranoside **62**. Displacement of the iodine with ammonium diethylthiophosphate in DMF gave the corresponding phosphorothioate **63**. One pot hydrolysis of the methyl-glycoside and acetylation of hydroxyl groups 1, 2, 3, was performed subjecting the methyl glycoside **63** to a 10:2.6:0.14 (v/v) mixture of acetic anhydride, acetic acid and H_2SO_4, affording compound **64**. Deprotection of ethyl groups was performed with TMSBr, in acetonitrile and pyridine as base, affording **65**. The remaining acetyl groups have been cleaved under Zèmplen conditions (with sodium methanoate in dry methanol), and after pH neutralization the thiophosphate **37** was obtained.



Scheme 11. Synthesis of thiophosphonate.

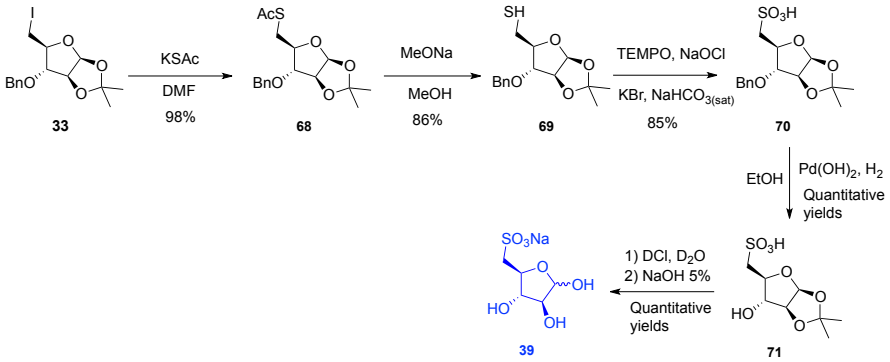
The sulphur-based anaolgues were synthesised starting from the intermediate **45**.

Using SO_3-Py complex as sulfation agent[140] in DMF, I synthesised the 5-sulfate compound **66**. The benzyl group was cleaved through $Pd(OH)_2$ catalysed hydrogenation affording **67** in quantitative yield. Final deprotection of the dimethylacetale in a D_2O solution and then pH neutralization, afforded the sulfate **38**.



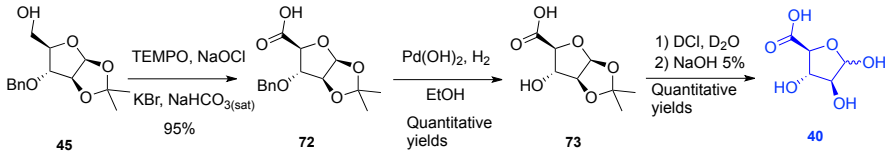
Scheme 12. Synthesis of sulphate analogue.

The sulphonate analogue **39** has been synthesised subjecting the 5-iodo compound **33** to a DMF solution of potassium thioacetate, affording **68**. Then the acetate group was removed by a sodium methanoate solution, giving the thiol **69** and the corresponding disulphide in low amount (<10%). The sulphur was oxidized to sulphonate **70** by sodium hypoclorite, using TEMPO and KBr as catalyzers in a sodium bicarbonate solution. The benzyl group was removed in EtOH by hydrogenation and $Pd(OH)_2$ as catalyser affording compound **71**. Deprotection of the isopropylidene was performed in a D_2O solution and after final pH neutralization sulphonate analogue **39** afforded in quantitative yield.



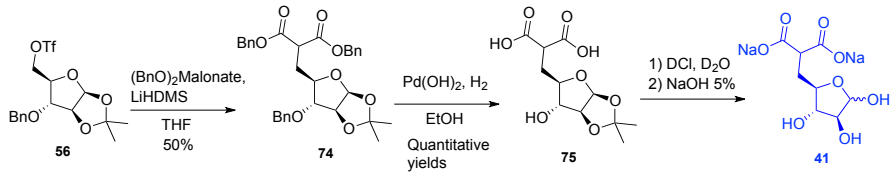
Scheme 13. Synthesis of sulphonate analogue.

Oxidation of the intermediate **45** with sodium hypoclorite, using TEMPO and KBr as catalyzers in a sodium bicarbonate solution[141], gave the corresponding carboxylate derivative **72**. Then the benzyl group was cleaved in EtOH, with $H_2/Pd(OH)_2$ in quantitative yield **73**. Final deprotection of the dimethyl acetal in a D_2O solution afforded the non isosteric carboxylate analogue **40** afforded after pH neutralization.



Scheme 14. Synthesis of carboxylate compound.

The malonate analogue **74** was synthesised reacting the freshly prepared triflate **56** with dibenzylmalonate [142] and LiHMDS as base, in dry THF. Then the three benzyl groups were cleaved in EtOH with $H_2/Pd(OH)_2$ affording compound **75** in quantitative yield. The cleavage of the last protective group dimethyacetal, was performed in a D_2O solution, and final pH neutralization gave the 5-malonate **41**. Unfortunately, the deprotect compound **41** was not stable and after lyophilisation ^{1}H-NMR showed the presence of secondary products, probably due to intramolecular lattonization with OH in position 3, as confirmed by mass analysis, that showed also the presence of (m/z = 217.1 [M-H_2O-H]$^-$).



Scheme 15. Synthesis of 5-malonate-D-arabinofuranose.

The ability of these A5P analogues (**31**, **34-41**) to bind API of *P. aureginosa* has been explored by STD-NMR experiments. This screening showed that only the isosteric phosphonate **31** and the difluoromethyl phosphonate **36** analogues are able to bind the enzyme in its catalytic site.

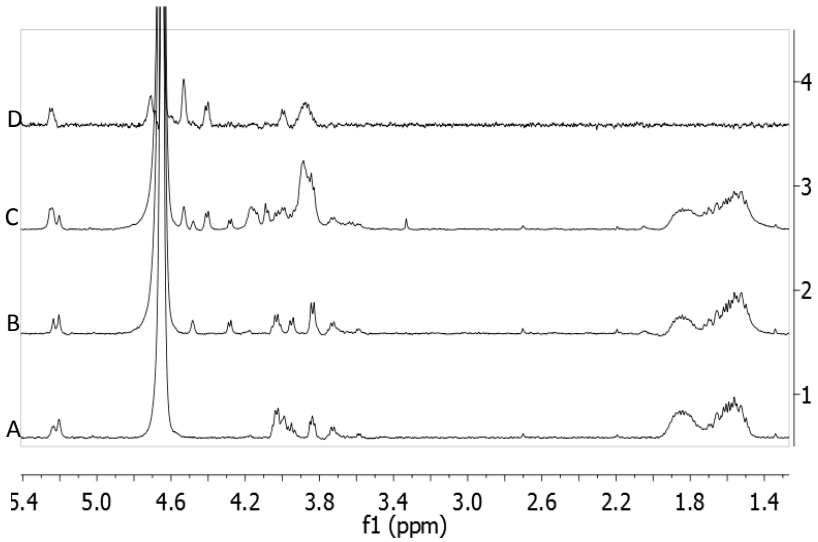


Fig. 10. A) ^{1}H NMR spectrum of a solution of compound **31**; B) ^{1}H NMR spectrum of a solution of compound **31** and KdsD; C) ^{1}H NMR spectrum of a solution of A5P, **31** (3 mM) and KdsD; D) STD-NMR spectrum of a solution of **31** (3 mM) and KdsD acquired with a saturation

time of 2.0 s and a saturation frequency of 7.9. All the spectra were recorded in phosphate buffer 50mM pH 7.0, at 37°C; compound-protein ratio was 100:1

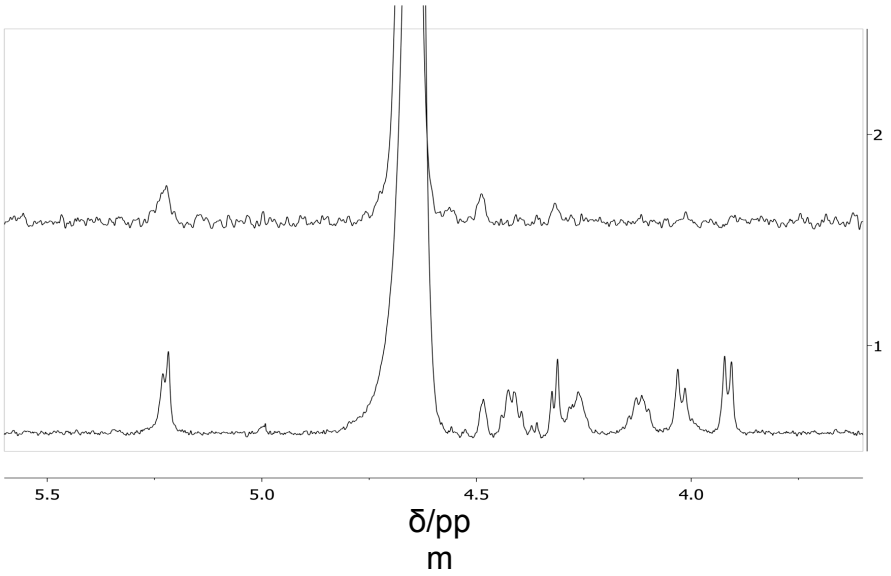


Fig. 11. A) ^{1}H NMR spectrum of a mixture containing **36** (3mM) and Pa-KdsD (30 uM); B). STD-NMR spectrum recorded on the same mixture with 3s of saturation time. The sample was dissolved in PBS, pH 7.4, 37°C

4.1. Akt: Introduction

The activation of survival pathways is a major feature of tumour cells, which contributes to the limited efficacy of antitumour treatment based on conventional cytotoxic drugs and on molecularly targeted agents. The deregulation of oncogenic signalling pathways providing survival advantages to tumour cells is mediated by several cellular networks that include different protein kinases. The human genome comprises more than 500 protein kinases including serine/threonine, tyrosine and dual specificity kinases.

Because of the important physiological and pathological role of kinases, the human kinome represents one of the most important drug discovery opportunities in different therapeutic areas, including oncology. Protein kinases comprise a large family of enzymes that catalyse the transfer of the terminal phosphate group from ATP (adenosine triphosphate) to protein substrates, specifically to the hydroxyl group of serine or threonine (Ser/Thr kinases) or tyrosine (Tyr kinases). Since protein kinases represent key players in many crucial cellular processes like proliferation, differentiation, and apoptosis, the discovery of small molecule kinase inhibitors has attracted growing interest for novel drug research and development, as well as the identification of experimental tools for the understanding of the biological roles of this class of proteins.

Among the kinases implicated in aberrant tumour cell signalling, the serine/threonine protein kinase B (PKB), also known as Akt, plays a key role as a component of the PI3K-Akt-mTOR axis.

Inappropriate activation of the PI3K/Akt pathway has been linked to the development of several human pathological states, including cancer[143]. The serine/threonine protein kinase B (PKB), also known as Akt, plays a key role as a component of the PI3K-Akt-mTOR axis in tumour cell survival. In fact, hyperactivation of Akt is a common feature of tumour cells that has been linked to increased survival, invasion and drug resistance. Activation of Akt occurs by a variety of mechanisms, including loss of PTEN, the phosphatase that counteracts PI3K-dependent Akt activation, mutations in the PI3K catalytic subunit, receptor tyrosine kinase activation and Ras activation. Increased Akt activity has been shown to induce oncogenic transformation of cells and tumour formation in the breast, prostate, ovary and pancreas. Due to its central role in these processes, Akt has recently received great attention as a promising molecular target in cancer therapy[144]. The available evidence supports that tumour cells with acquired resistance to antitumour agents may display increased Akt activation[145]. Also, treatment with molecularly targeted agents (e.g., mTOR inhibitors) can result in activation of feedback loops involving Akt[146]. Alterations in sialidase NEU3 mostly influence or lead to a malignant phenotype, including uncontrolled growth, invasion and metastasis: NEU3 activated by IL-6 directs IL-6-mediated signalling via the PI3K/Akt cascade in a

positive feedback manner and thus contributes to a malignant phenotype, including suppression of apoptosis and promotion of cell motility in renal cell carcinoma ACHN cells[147]. Three members of the Akt family (Akt1, Akt2 and Akt3) have been identified which are all structurally composed of an N-terminal pleckstrin homology (PH) domain, a central kinase catalytic domain and a C-terminal hydrophobic regulatory motif (HM)[148]. The three Akt isoforms are ubiquitous, Akt1 being the dominant one. Akt2 has been implicated in insulin signalling, whereas Akt3 is mainly expressed in brain, skeletal muscle and liver.

Akt has been shown to be activated in a phosphatidylinositol 3-kinase (PI3K)-dependent manner: PI3K phosphorylates the 3-position of phosphatidylinositols giving rise to three signalling phospholipids (PI(3)P's). Binding of the PH domain of Akt to membrane PI(3)P's is fundamental for Akt activation, causing a conformational change in the protein and its translocation to the plasma membrane; phosphorylation of Thr 308 of the activation loop in the kinase domain and Ser473 of the hydrophobic motif (HM) completes Akt activation.

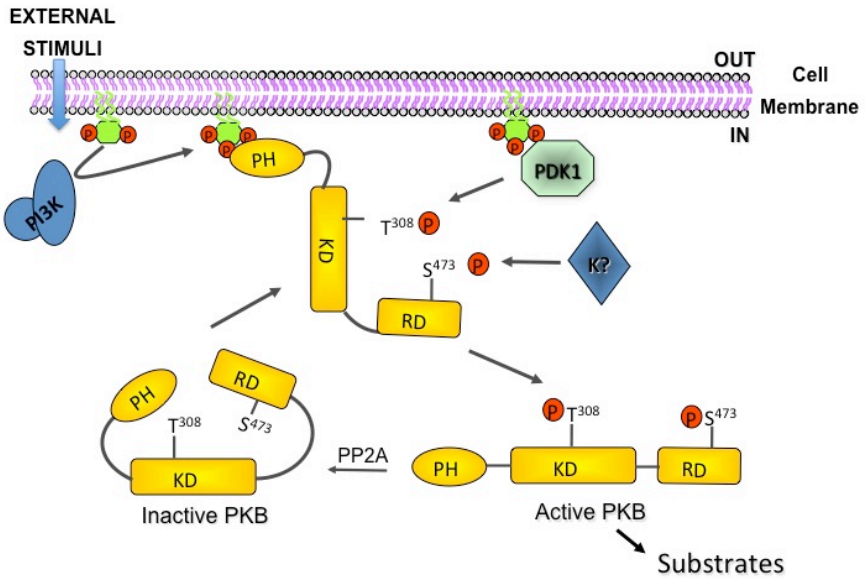


Fig. 12. Schematic rapresentation of Akt signaling patway.

In light of these considerations, intense efforts towards the synthesis of Akt inhibitors have recently emerged. Part of these drug-development efforts has been focused on ATP-binding-site inhibitors (Fig. 13). However, since the ATP-binding-site is highly conserved in kinases, the design of selective inhibitors is a formidable challenge. A second valuable approach to Akt-specific inhibitors is the development of substrate–competitive inhibitors; because the substrate-binding domain is less conserved than the ATP pocket, they should be more specific[149]. Allosteric inhibitors have also been proposed[149]. Akt inhibitors tested in clinical trials have shown adverse effects associated with alteration of insulin/glucose metabolism. Thus, there is the need to generate novel compounds

with improved pharmacological profiles. An additional promising approach is to target the pleckstrin homology (PH) domain; over the past few years a few inhibitors based on this approach have been developed[149]. Unlike the ATP-binding domain, which is present in all kinases, the PH domain is present in a relatively small number of signalling kinases, suggesting that inhibition of the PH domain of Akt might achieve better selectivity than the ATP-pocket-binders[149].

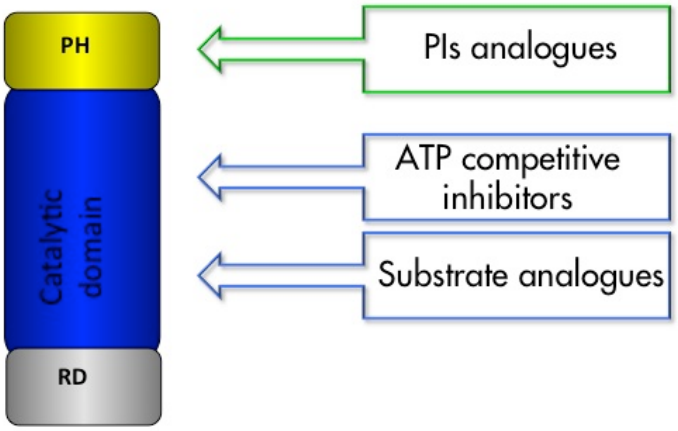


Fig. 13. Schematic rapresentation of Akt and its possible targets for drug discovery.

Several lines of evidence support the relevance of Akt in maintaining survival of cancer cells. It is now clear that antitumour therapy based on a single molecularly targeted agent does not result in the cure of patients, but at best in increased progression free survival. Thus, the development of drug combinations including Akt inhibitors is promising in an attempt to increase the efficacy of clinically

available antitumour agents. Finally, the inhibitors developed in the context of the project may be useful tools to improve the understanding of the redundancies of signalling pathways in cancer cells.

Hence, the idea is to use the inhibition of Akt by novel compounds targeting the PH domain to define the contribution of Akt to tumour biology with special emphasis on chemoresistance and invasiveness, using cancer cell lines of different tumour types, including ovarian and lung carcinoma cells resistant to chemotherapeutic agents in clinical use and thyroid cancer cell lines displaying constitutive Akt activation sustained by oncogene activation. Based on results on target inhibition and on modulation of tumour cell behaviour after treatment, other compounds will be designed and synthesized as new PH-Akt inhibitors, possibly endowed with improved pharmacological profiles. Additionally, the obtained data will also be helpful to increase the comprehension of anti-cancer drug resistance mechanisms.

The new potential kinase inhibitors will be synthesized using glucose as a scaffold to mimic the inositol structure.

The key features needed for interaction with the PH domain are: i) an anionic group mimicking the phosphate group in position 3 of the inositol (position 6 of the sugar), which should be able to form a salt bridge with a key arginine residue[150]; ii) a β-configured substituent at the anomeric carbon of D-glucose (position 1 of inositol) possessing a hydrogen-donor heteroatom and a lipophilic

group mimicking the diacyl glycerol moiety of phosphoinositides. However, the long lipophilic acyl chains of the natural substrate can be substituted by smaller hydrophobic groups, since the long fatty acid chain is not needed for enzyme recognition, but only for membrane anchoring[151]. In addition, the lack of the hydroxyl group in position 2 of the inositol ring, substituted by the pyranosidic oxygen of the sugar, does not alter enzyme recognition.

4.2. Akt: Results and discussion

The idea is to design and synthesise new potential kinase inhibitors using glucose as a scaffold (Fig. 14) in order to mimic the inositol structure.

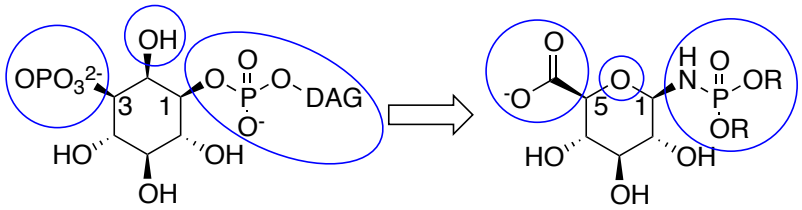


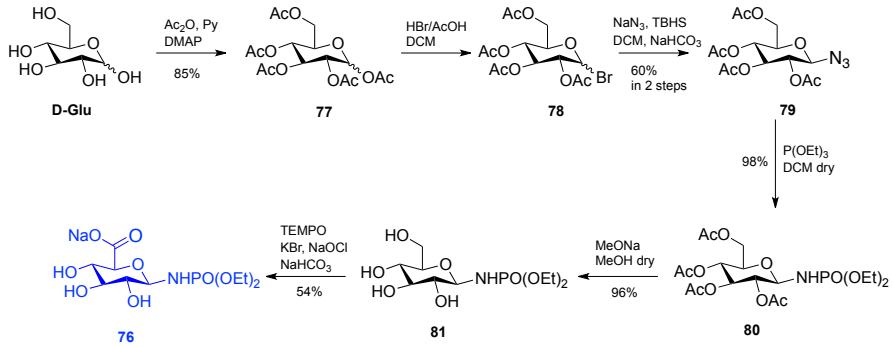
Fig. 14. Comparison between phosphatidylinositols and the designed glucose-scaffold analogue.

The phosphate in position 1 of the inositol has the same configuration of the anomeric substituent in β-D-glucose; the exocyclic hydroxyl group in position 2 is mimicked by the endocyclic oxygen of the glucose scaffold, and the carbon in position 3 of the inositol ring has the same configuration of the glucose carbon in position 5. Hence, β-D-glucose could be used as a scaffold for the synthesis of potential analogues of the initsitol ring. The glucose derivatived that I synthesised (compound 76, Scheme 16) has the following characteristics: the metabolically labile phosphate ester linkage between the inositol ring and the diacylglycerol moiety (DAG) was substituted by a phosphoramidate group as a stable phosphate mimic. The phosphate group in

position 3 of the inositol ring (position 5 of the sugar) was substituted with the carboxylic group. As the degradation of inositol phospholipids occurs through the enzymatic action of PI-specific phospholipase C (PIPLC) with the formation of a 1,2 cyclic phosphate, 19 deletions of the inositol 2-OH group to block the formation of the cyclic phosphate is an important characteristic of the synthesized mimetic, and carbon 2 was replaced by the endocyclic oxygen of the glucopyranose ring. Finally, lipophilic acyl chains of the diacyl glycerol moiety were mimicked by two ethyl groups on the phosphoramidate, reasoning that the long chains are not needed for enzyme recognition, but only for membrane anchoring.[151]

D-Glucose was peracetylated in a pyridine solution of acetic anhydride, with DMAP as catalyse (Scheme 16). Reaction of 1,2,3,4,6-penta-O-acetyl-D-glucose **77** with Bromidric acid/Acetic acid in dicloromethane afforded 2,3,4,6-tetra-O-acetyl-D-glucosyl bromide (**78**). The β-D-glucopyranosyl azide **79** was prepared by the reaction of the freshly prepared glucosyl bromide **78** under phase transfer catalysis with tetrabutylammonium hydrogen sulfate (TBAHS) and sodium azide.[152] Alkyl phosphoramidates **80** was synthesized by Staudinger reaction of azide **79** with the corresponding trialkyl phosphite, as reported by Kannan and co-workers who proposed the synthesis of several glycosyl phosphoramidates as isosteric analogues of native glycosyl phosphates.[153] The product was obtained with total stereoselection

at the anomeric position as determined by ^{1}H NMR. The coupling constants between H-1 and H-2 (around 9.0 Hz), typical for a trans-diaxial arrangement of the substituents, indicated a b-orientation of the azido group in a $^{4}C_{1}$ chair conformation. ^{31}P NMR of derivatives 2–7 confirmed the presence of the NH–P=O group in which the phosphorous atom resonates at approximately 8–10 ppm. Deprotection of phosphoramidate **80** under Zèmplen conditions (cat. Na, MeOH) afforded compound **81**, which was selectively oxidised at the primary hydroxyl group by 2,2,6,6-tetramethyl-1-piperidinyloxyl/hypochlorite (TEMPO), [154] to the corresponding uronate (**76**).



Scheme 16. Synthesis of β-D-Glucuronyl diethyl phosphoramidate.

Glucose derivatived **76** and other compounds previously synthesised in our lab [150], were assayed for in vitro inhibitory activity against Akt1, using an in vitro Kinase Assay Kit (StressXpress_non-radioactive Akt/PKB Kinase activity), Fig. 15. These preliminary data show that phosphoramidate **76** is the most active derivative

among the synthesized compounds, followed by glucose diethyl phosphoramidate **I** (Fig. 15); it is worth noting that uronic acid **76** is as active as the parent glucose derivative **I**. It can also be noted that the uronic acids are more active than the corresponding glucose counterparts, suggesting that the carboxylic group may be relevant for interactions of the putative inhibitors with the binding site.

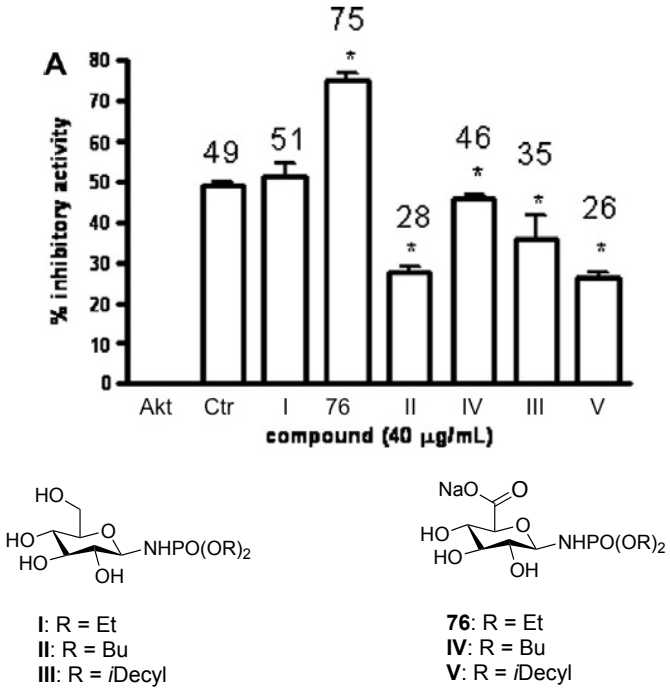


Fig. 15. In vitro Akt Kinase Assay (ELISA) of compounds **76**, **I-V**, 1L6hydroxymethyl-chiro-inositol-2-(R)-2-Omethyl-3-O-octadec-yl-sn-glycerocarbonate was used as the positive control (CTR). Statistical analysis was done using the one-way ANOVA followed by Tukey's test as post hoc comparison; *P <0.05 versus CTR;

With the aim of disclosing key atomic details related to the binding of the synthesised phosphatidylinositol-3-phosphate analogues, a set

of docking simulations was carried out on the most active compounds **I** and **76**. The adopted docking protocol [150], which was already used for the investigation of ligand–receptor interactions,[155] was initially validated on the X-ray structure of the complex between Akt PH domain and PtdIns(3,4,5)P3. The relative stability data obtained by molecular modelling studies on compounds **I** and **76** are summarized in Fig. 16.

Ligands	Relative stability
76	0
I	+16.3

Fig. 16. Relative stability data (kcal/mol), which correspond to ($E_{docking\ Ligand}$ - $E_{docking\ 76}$) related to the binding of the inhibitors **76** and **I** to the PH domain.

According to this computational analysis, binding of the uronic acid **76** is favoured compared to the corresponding glucose derivative **I** (Fig. 17), in nice agreement with preliminary in vitro assays. Indeed, both compounds bind to Akt in a similar fashion (Fig. 17) with the glucose OH groups involved in H-bond interactions with Lys14, Arg25 and Arg86 of the pleckstrin homology domain of Akt. Moreover, the phosphoramidate group interacts, via H-bond, with Glu17. The higher affinity of **76** can be essentially ascribed to the chelating interaction between the carboxylic group and Arg23 (Fig. 17B), which is weaker and partially lost in the $–CH_2OH$ analogues (Fig. 17A). It is also important to note that, on the basis of the

docking simulations, the length of the alkyl chains in the phosphoramidate group has no relevant effects on binding, as already reported by Kozikowski and co-workers.[156] These results are in agreement with molecular modelling studies of the natural substrates and inositol mimics already reported.[156]

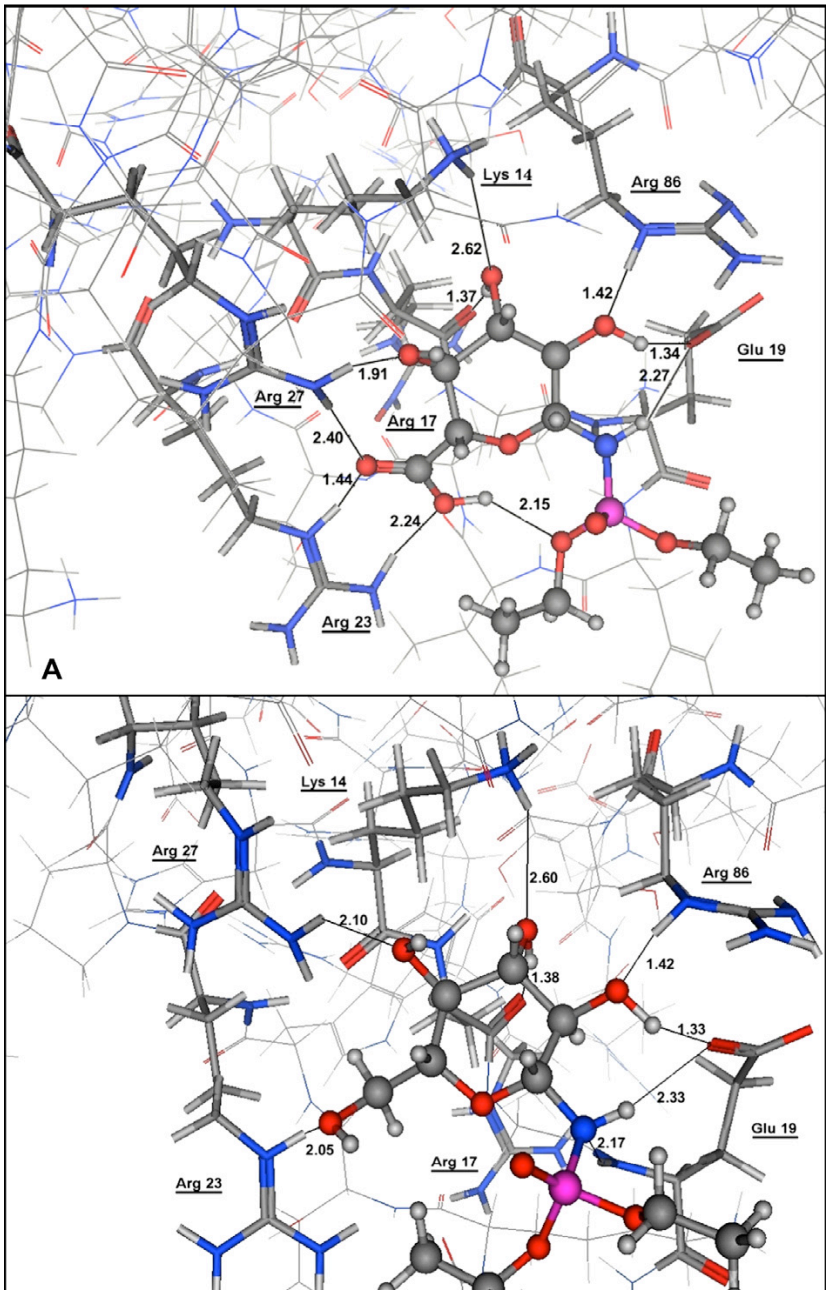


Fig. 17. 3D view of the binding mode of **I** (A) and **76** (B) to the PH domain. Relevant interatomic distances are shown in Å.

ELISA tests and molecular modelling studies highlighted compound **76** as a good hit as an Akt inhibitor. Hence, in vitro activity on dendritic cells (DCs) as a cellular system was also carried out, since it has been recently reported that Akt1 has an essential role in DC function.[157] DCs are a special class of leucocytes that play a fundamental role in regulating innate and adaptive immune responses.[158] They can be activated following encounters with microbial cell products that signal through specific receptors called toll-like receptors (TLRs).[159] TLR engagement leads to nuclear factor (NF)-kB activation through the MyD88-dependent pathway[160] and also through the activation of the phosphatidylinositol-3-kinase (PI3K) that phosphorylates its downstream target Akt. Inhibition of Akt phosphorylation leads to partial DC activation and reduced DC functionality.[161] To test the inhibitory activity of compound **76** we took advantage of the well-characterized murine DC line D,[162] which can undergo maturation in vitro upon encountering of microbial cell products, such as lipopolysaccharides (LPS), leading to the activation of PI3K/Akt pathway, required to have efficient IL-2 production.[163] The ability of compound **76** to inhibit LPS-induced Akt phosphorylation in D1 cells was investigated by both Western blot and cytofluorimetric analysis. In particular, D1 cells were pretreated with compound 5 for 1 h and then incubated with LPS (1 μg/mL) for an additional 20 min. Immunoblotting experiments (Fig. 18) were performed using a Ser473 phospho-specific antibod.

Compound **76** inhibited Akt phosphorylation in a dose-dependent manner (IC_{50} 97 μM).

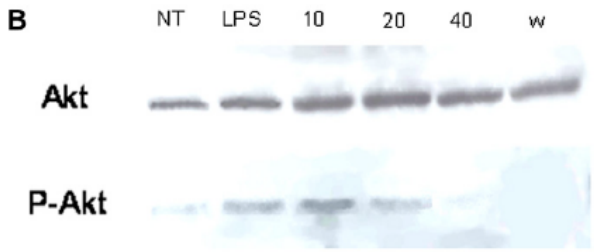


Fig. 18. Western blot analysis showing the inhibitory activity of compound **76** with D1 cells at the indicated concentrations (μg/mL); Wortmannin (W) was used as the positive control. The protein levels of phosphorylated Akt (P-Akt) and Akt were determined using monoclonal antibodies anti-phospho-Ser473-Akt.

To assess whether compound **76** could be cytotoxic, toxicity was investigated with the annexin V test. D1 cells did not show any appreciable percent of apoptotic cells after 24-h incubation with LPS and compound **76** at the concentration showing maximal inhibitory activity. Further biological investigations were carried out to best characterize the biological activity of compound **76**. In particular, inhibition of cell maturation and cytokine production induced by LPS was analysed. To this end, D1 cells were pre-treated with increasing amounts of the inhibitor **76** for 1 h and then cultured in the presence of LPS (1 μg/mL) for the remaining 24 h. D1 cell maturation state was analysed by evaluating the up-regulation of the co-stimulatory molecule B7.2 and the ability of D1 cells to produce inflammatory cytokines was tested by evaluating IL-2 concentration

in the supernatant. Wortmannin was always used as the positive control. Similarly to wortmannin, compound **76** did not block the up-regulation of B7.2 induced by LPS and inhibited IL-2 production in a dose-dependent manner (Fig. 19).

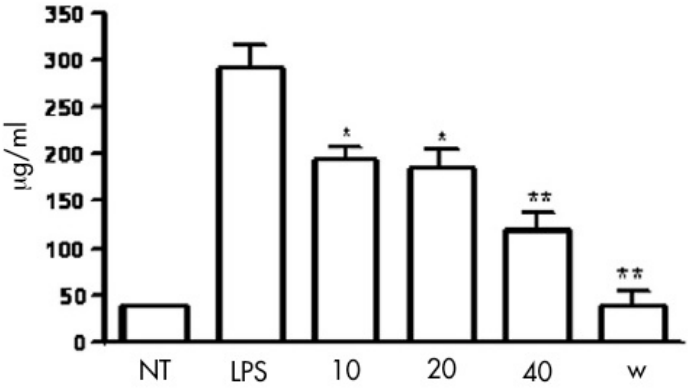


Fig. 19. IL-2 levels in the supernatant of D1 cells cultured for 24 h in the presence of LPS with the indicated amounts (µg/mL) of compound **76**. Wortmannin (W) was used as the positive control. Statistical analysis was done using the Kruskal–Wallis ANOVA for non-parametric values followed by Dunns test as post hoc comparison. **P <0.001 versus LPS, *P <0.05 versus LPS.

Recent studies showed that PI3K/Akt pathway is also involved in cardiomyocytes function.[164] To obtain preliminary information on the cardiac effect of inhibitor **76**, we studied its effect on contraction (unloaded twitch) of adult rat ventricular myocytes during field stimulation at 2 Hz after silencing the expression of Akt1, the main kinase activated through PH-domain interaction, by RNA interference. As shown in (Fig. 20), compound **76** effect was completely abolished in Akt1-silenced myocytes, thus suggesting that

it indeed occurred through PH-domain antagonism and, at the same time, proving the involvement of Akt signalling pathway. The positive effect of compound **76** on cardiomyocyte contractility was unexpected and suggests the absence of acute cellular toxicity. Nevertheless, depending on their nature, cellular mechanisms responsible for acute inotropy may lead to cell damage if sustained chronically (e.g., cell Ca^{2+} overload). Therefore, further experiments will be required to assess the mechanism of compound **76** inotropic effects and rule out that it may result in myocardial damage during long-term administration.

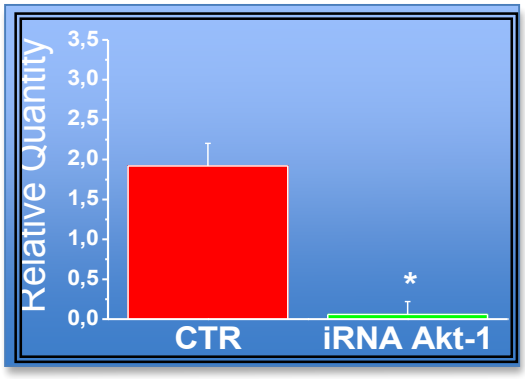


Fig. 20. Effects of compound **76** on twitch amplitude on not-transfected (CTR) and Akt1-silenced cardiomyocytes (iRNA Akt-1).

In addition aberrant activation of the PI3K/Akt (PKB) pathway plays a fundamental role in the tumourigenesis and progression of thyroid cancers [165]. The major oncogenic mutations found in papillary thyroid carcinomas (PTCs) involving BRAF, RAS, PI3KCA, RET genes have been demonstrated to activate the PI3K/Akt pathway. Thus,

targeting this deregulated signalling cascade represents an attractive therapeutic approach for treatment of PTC tumours. Moreover, it has been reported that thyroid cancer cells characterized by PI3K/Akt signalling deregulation show a preferential sensitivity to inhibitors in this pathway[166]. Thus sensitivity of tumour cells to the studied compounds was examined on the human ovarian carcinoma cell line IGROV-1 using growth-inhibiton assays. A 24h exposure to the compounds indicated that **76** displayed marginal inhibitory effects when tested at 50 μM. (Table 2).

Compound	IC_{50} [μM][b]
	IGROV-1
76	> 50

Table 2. [a]Cell sensitivity of the human ovarian carcinoma IGROV-1 cells to the drug was assessed by growth-inhibition assay. 24 h after seeding, cells were exposed to the compounds for 24 h and were counted 48 h later.
[b]IC_{50} = drug concentration inhibiting growth by 50%.

5.1. CI-M6PR: Introduction.

Mannose-6-phospate (M6P) acts as a trafficking marker for delivery of enzymes to the lisososome[167].

M6P bearing proteins are transported in vescicles to the Golgi apparatus (Fig. 21) where they can interact with two receptors: the cation-dependent (46kDa) Ma6P receptor (CD-M6P), and the 300 kDa cation-independent receptor (CI-M6PR). CD-M6PR and CI-M6PR are the only two members of the P-type lectins family, and sequence comparisons between them shows that they are related and that they diverged from a common ancestral gene.

The 46 kDa CD-M6P receptor is a single peptide chain and is known as cation-dependent because of its ability to bind the ligand in presence of divalent cation.

The CI-M6P receptor is also known as the M6P/IGF-II receptor, as it has been shown to bind insulin-like growth factor II (IGF-II)[168] and retinoic acid[169]. IGF-II is the best characterized non M6P-bearing ligand of CI-M6PR. This hormone plays key roles in metabolic regulation and foetal growth through three types of receptors, two tyrosine kinase receptors (IGF-I receptor and insulin receptor isoform A) and CI-M6PR. The stimulation of the type 1 receptor tyrosine kinase induces a protein phophorylation cascade leading to biological effects, particularly the insulin-mediated growth and increase in CI-M6PR expression[170], which mediates endocytosis and clearance of IGF-II[171]. The lysosomal degradation of IGF-II is critical

because elevated levels of IGF-II induce overgrowth. Though CI-M6PR circulates at a concentration lower than that of IGF-II, it seems to play a role in IGF-II clearance. Besides IGF-II, retinoic acid and lysosomal enzymes carrying M6P, plasma membrane CIM6PR has been shown to bind a number of other extracellular ligands bearing the M6P recognition marker.

Accumulation of the protein-receptor complexes leads to the formation of clathrin-coated vesicles that can travel from the Golgi apparatus to a pre-lysosomal compartment. Here, a pH drop causes dissociation of the complex (pH less than 5.5 is required). After dissociation the M6P bearing proteins are delivered to the lysosome, while the receptors are recycled into the Golgi apparatus. However, the CI-M6PR, which participates in this cellular routing, is also anchored to the cell surface membrane and can internalize extracellular ligands[172]. In summary, the pool of lysosomal acid hydrolases comes both from in-situ synthesis, which involves the two receptors, and from endocytosis through the CI-M6PR[173].

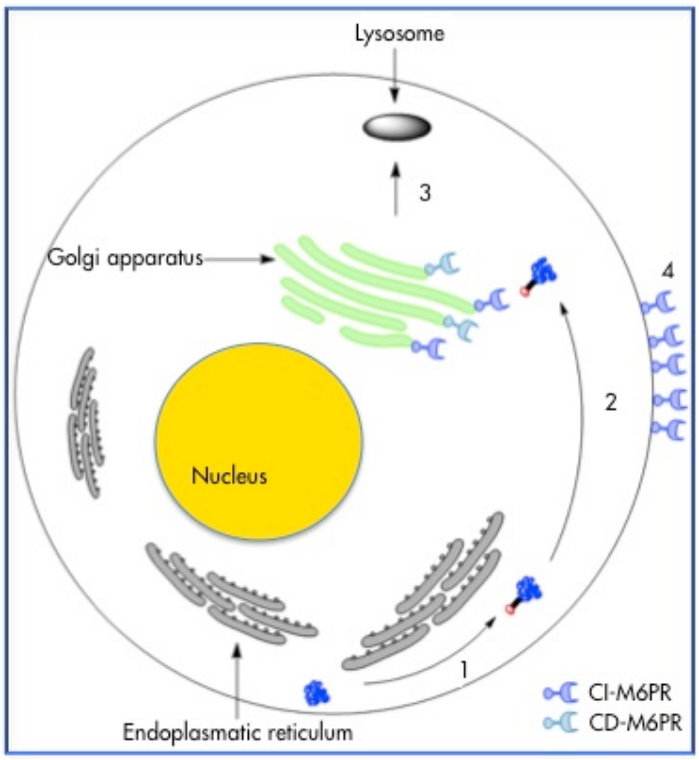


Fig. 21. Schematic cellular routing of M6P bearing ligands. 1) At the endoplasmic reticulum, enzymes obtain the Man-6-P ligand; 2) Transport of enzymes to the Golgi apparatus in vesicles; 3) At the Golgi apparatus, Man-6-P bearing enzymes interact with the M6P/IGF-II and CD-M6P receptor and are taken to the lysosome; 4) The M6P/IGF-II receptor present on the cell surface allows endocytosis of extracellular Man-6-P bearing proteins.

MPRs have a relatively long half-life ($t_{1/2}$=20h), and both receptors display optimal ligand binding at pH 6.5 and no detectable binding at pH 5.

Only 10% of the MPRs are found at the cell surface, while the remainders of the receptors are found predominantly in endosomal compartments[174].

According to the type of ligand, CI-M6PR may modulate a large panel of biological pathways, such as cell migration, wound healing, angiogenesis, cell growth inhibition, apoptosis or viral infection[175]. Therefore, CI-M6PR constitutes a target for several clinical applications, which justifies the development of synthetic M6P analogues to compensate the poor stability of natural M6P in human serum and increases the binding stability with the receptor.

CI-M6PR

The extracytoplasmatic region of the CI-MPR has a repetitive structure of 15 domains and contains three Man-6-P pocket bindings[176]. Expression of recombinant truncated forms of the CI-MPR has mapped its three carbohydrate binding sites: two high affinity sites ($Ki \approx 10$ μM for Man-6-P) map to domains 1–3 and domain 9, while domain 5 houses a low affinity ($Ki \approx 5$ mM for Man-6-P) binding site[176]. A comparison of the binding properties of the individual carbohydrate recognition sites demonstrated that domain 9 of the CI-MPR exhibits optimal binding at pH 6.4–6.5, similar to that of the CD-MPR. In contrast, the N-terminal Man-6-P binding site (i.e., domains 1–3) has a significantly higher optimal binding pH of 6.9–7.012. This observation may not only explain the relatively broad pH range of ligand binding by the CI-MPR, but also the ability of the CI-MPR, as opposed to the CD-MPR, to internalize exogenous ligands at the slightly alkaline pH 7.4 present

at the cell surface. Domain 9 of the CI-MPR, like the CD-MPR, is highly specific for phosphomonoesters, while recent studies demonstrate that domain 5 of the CI-MPR exhibits a 14- to 18-fold higher affinity for Man-P-GlcNAc than Man-6-P[13] The presence of three distinct carbohydrate recognition sites in the CI-MPR, likely accounts for the ability of the CI-MPR to recognize a greater diversity of ligands than the CD-MPR.

The crystal structures of one of the Man-6-P binding sites (domains 1–3) and the IGF-II binding site (domains11–14) have been obtained [174].

The CI-MPR seems to be a dimer in the membrane, although it behaves like a monomer in detergent solutions under most circumstances. Receptor dimerization allows for high affinity binding of ligands that are multivalent for M6P residues[177]. What is not clear is how the various ligand binding sites are oriented relative to each other and whether the CI-MPR undergoes conformational changes that may be influenced by pH or ligand binding.

Kornfeld's early studies[178] involved isolation of native mixtures of mannose-rich oligosaccharides (Fig. 22) from cellular glycoproteins. These studies revealed that most Asn-linked oligosaccharides carry one or two phosphorylated mannoses, with the latter binding more tightly than the former.

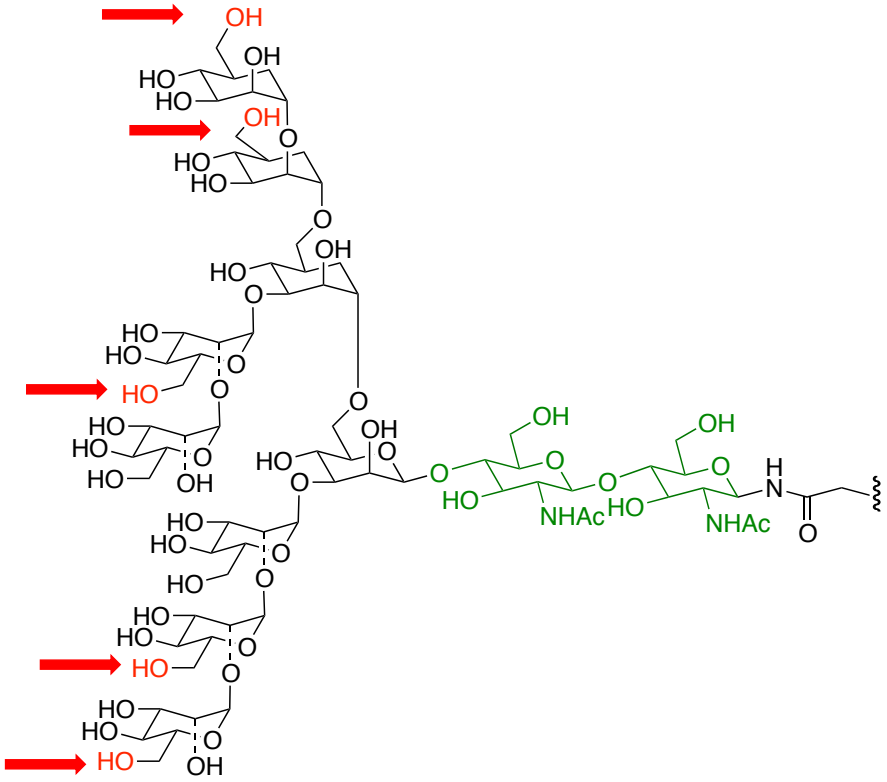


Fig. 22. High mannose-type N-linked oligosaccharide. The phosphorylation sites are enlightened in red.

Later, elegant work by Hindsgaul revealed that one could gain nearly an order of magnitude in binding to the CI-MPR with pentasaccharides (Fig. 23) bearing two M6P residues, as opposed to one[179]. However, neither of these studies could reproduce the high relative binding affinities (RBAs g vs M6P) seen with native ligands such as hGUS (human β-glucuronidase) or the synthetic glycoprotein, PMP-BSA (pentamannose phosphate-bovine serum albumin).

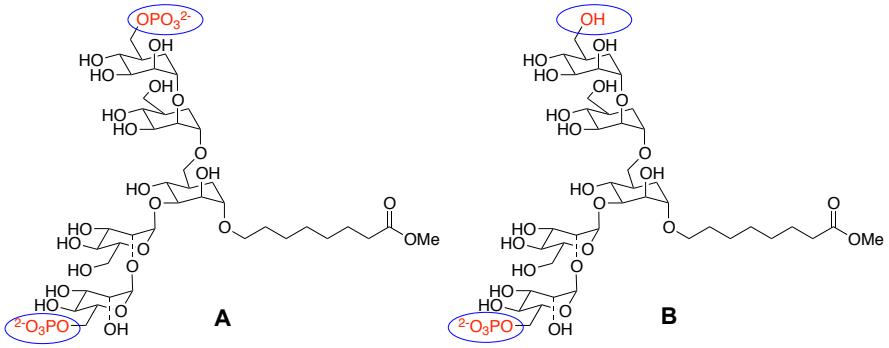


Fig. 23. Monophosphate **B** displays 10-fold lower M6P/IGF2R binding affinity than the bis(phosphate) ligand **A**.

β-glucuronidase (hGUS), a homotetrameric lysosomal enzyme bearing multiple M6P moieties, stabilized the receptor's dimeric structure by cross-bridging the M6P binding sites on two adjacent subunits[180]. These data support a dimeric model for binding of bivalent M6P-based ligands by the M6P/IGF2R. Importantly, they also observed that hGUS binding increased the rate of internalization of the receptor and consequently stimulated the degradation of any passenger ligands, including IGFII, by 3- to 4-fold.

Multivalent interactions between receptors and their ligands, which are common in biology, involve a multistep mechanism in which most of the entropic cost is paid by the initial binding event, and subsequent contacts contribute a favorable enthalpy without further sacrifice of rotational and translational entropy[181]. The resultant high binding affinity in these interactions is due to a reduced rate of ligand-receptor dissociation. This type of interaction occurs in carbohydrate binding to lectins and is particularly important in the

binding of M6P-bearing oligosaccharides by P-type lectins, such as the M6P/IGF2R.

Given that (at least) two M6P-binding pockets are available per receptor in the monomeric binding model and four per receptor in the dimeric model, bi- or multivalency may account for this effect. Two models of bivalent binding have been hypothesized, a first where only the monomeric unit of the receptor is responsible for the divalent binding (Fig. 24A) and a second one involving the dimeric receptor (Fig. 24B). However, experiments with mutations in domain 3 and its consequent inactivation [178], confirm model B as the most realistic one.

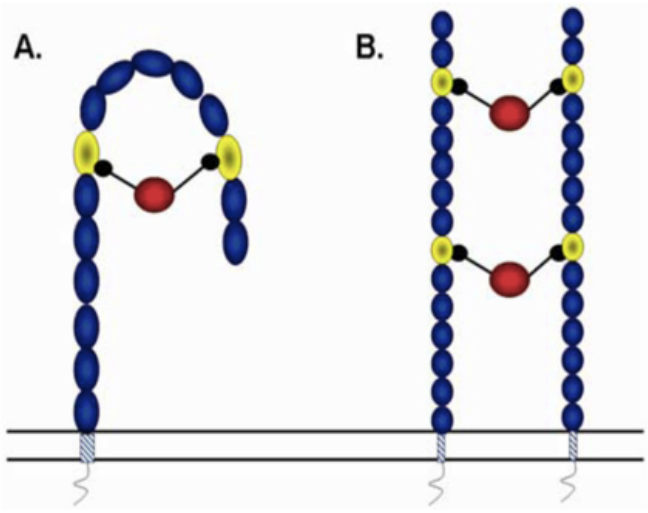


Fig. 24. Two models of bivalent binding hypothesized: A) only the monomeric unit of the receptor is responsible of the divalent binding. B) involving the dimeric receptor.

The intermolecular distance between domain 3 and 9 is estimated to be around 45 Å, while the interphosphate distance between the two M6Ps caps of a bis-phosphorilated oligosaccharide should be from 26 to 30 Å[182]. However, these data are from preliminary modeling studies based on the crystal structure of domains 1–3 using topographical information based on the amino acid sequence of each domain.

Different studies [182,142] have been done to find and prove the correct interphosphate distance, but unsuccessfully, maybe because they used very flexible bi-anthenary compounds to determine the dimeric binding, so the distances calculated could not fit with the real one.

Bhock[183], using a peptide scaffold, synthesized a compound with high affinity for the receptor, but this was not because of a divalent binding, but thanks to hydrophobic interactions between a receptor pocket and an aromatic residue on the peptide scaffold.

So the finding of a molecular ruler that allows the divalent binding with CI-M6PR is still a very important challenge.

Some structural features are known to be crucial for the binding of M6P to CI-M6PR:

The hydroxyl group in position 2 of the piranose ring must be axial to allow a strong binding to the receptor[184].

Structural modifications to the anomeric position do not impair the binding to CI-M6PR, since fructose phosphate is recognized as easy as M6P [184].

The distance between the negative charge and the piranose ring is very important and has to be in the same range as in M6P, since the isosteric Man 6-methylphosphonate is well recognized, while the non isosteric Man 6-phosphonate displays a weak affinity[185].

A single negative charge is sufficient to allow the binding to the receptor, and a phosphorous atom is not strictly needed[175]; however, the best ligands contain two negative charges and are isosteric of M6P.

Linear mannose sequences that contain a terminal M6P linked α(1-2) to the penultimate Man were shown to be the most potent inhibitors[185].

Because of the biological relevance of M6PR, this receptor can be an important target for different diseases. Here I just mention a few examples.

Enzyme replacement therapy

The importance of this phosphomannosyl recognition system in the biogenesis of lysosomes is illustrated by the existence of over 40 different human lysosomal storage diseases that are estimated to affect 1 in 5,000 live births.

Lysosomal storage diseases (LSDs) are a class of genetic disorders in which lysosomal degradation is incomplete, due to the abnormal functioning of specific lysosomal enzymes.

The phenotypic outcomes include physical or neurological abnormalities with patients often having a reduced life expectancy.

A therapeutic strategy for the treatment of LSDs is the external administration of the deficient enzyme. This is known as enzyme replacement therapy (ERT). A strategy employing Man-6-P modified replacement enzymes would be appropriate, since the M6P/IGF-II receptor can allow for endocytosis of Man-6-P bearing extracellular proteins. The potential of ERT using the M6P/IGF-II receptor has been demonstrated for several enzymes, many of which are now used in clinical treatment[186]. Preparations of enzymes containing more Man-6-P residues have shown greater uptake into the lysosome[187] and in a wider range of tissue types than standard enzyme preparations. This highlights the usefulness of Man-6-P as an uptake enhancer.

However, new M6P analogues are needed to decrease the high treatment costs: these analogues have to be simpler than the oligosaccaride bearing M6P; they have to improve the affinity with the receptor and their stability in the blood circulation.

Cancer therapy

Even though the CI-M6PR expression is decreased in some malignancies such as hepatocarcinoma, its over-expression in the majority of solid tumours, particularly in breast cancers[188], indicate that this receptor could be considered as a way to address cytotoxic drugs to lysosomes. The routing of cytotoxic drugs *via* CI-M6PR may induce the lysis of lysosomes and then cell death. The higher specificity of CI-M6PR targeting in cancer cells versus normal cells

could be due to the higher expression of CI-M6PR and its higher affinity at slightly acidic pH (pH 6–6.5), as found in the extracellular environment of solid tumours. This approach could be complementary to the classical chemotherapy. In fact, most anti-neoplastic drugs, such as anthracyclines (*e.g.*, doxorubicin) are weakly basic molecules which target the neutral pH compartment of tumours, but become inactive in a more acidic environment.

Thus, the M6P analogues could also be used to target drugs to lysosomes where they would be activated by the large panel of acid protease activities of these structures. This approach could be used in the case of photo-sensitive, or pH-sensitive, drugs for imaging or therapy in cancers that express high levels of CI-M6PR[175].

Liver fibrosis

Liver fibrosis is a public healthcare threat in both developing and western countries, with 180 million people affected worldwide.

The key factor in the pathogenesis of liver fibrosis is the activation and proliferation of hepatic stellate cells (HSC) and their transformation into myofibroblasts. HSC contribute largely to the intrahepatic connective tissue expansion during fibrogenesis. Therefore, this cell type is an important target for antifibrotic therapy. Binding sites expressed on (activated) HSC were considered for their ability to serve as potential targets for carrier molecules. One of these is the mannose 6-phosphate/insulin like growth factor II

(M6P/IGF-II) receptor, whose expression is increased on activated HSCs, particularly during fibrosis.

Different studies have already shown that it is possible to selectively target HSCs by CI-M6PR, modifying proteins with M6P residues.

Human serum albumine modified with M6P with different degrees of modification was tested as a selective carrier for hepatic stellate cells (HSC); it was found that high M6P density ($M6P_{26}$-HSA) permits the selective target of HSC, while a lower number of M6P bring to a non-selective uptake from kidney and hepatic cells[189]. $M6P_{26}$-HSA has been also used to deliver glycyrrhetinic acid to HSCs[190].

Bovine serum albumine modified with M6P ($M6P_{20}$-BSA) and coupled with a triplex-forming oligonucleotide, targets specifically HSCs in rats[191].

The biological importance of the CI-M6P receptor is evident; however, at least two important challenges have to be addressed: the finding of new stable M6P analogues with high affinity for the receptor and the finding of the correct interphosphate distance to promote the divalent binding.

Hence, I chose to design and synthesise a new fluorinated analogue of M6P and work on a new concept of molecular ruler in order to discover the right distance able to stimulate the divalent binding.

5.2. CI-M6PR: Results and discussion

Until now a few mimetics of M6P have been synthesized[175].

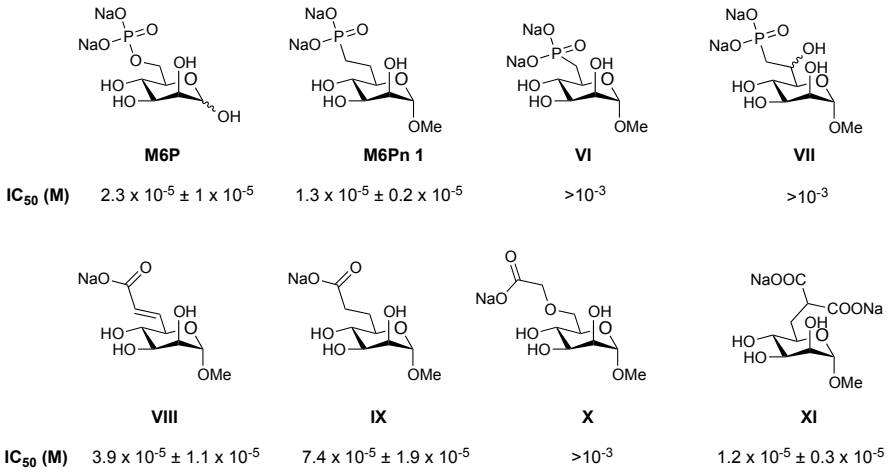


Fig. 25. Set of mannose-6-phosphate analogues and IC_{50} for M6PR.

As I previously said, the first idea is to synthesize a new M6P analogue, with a methyldifluoro-phosphonate group instead of a phosphate.

New analogues that are more stable than the phosphate group are required [175] and $CF_2PO(OH)_2$ could be a new promising solution. Phosphatase enzymes cannot hydrolyze it, so we can extend the circulation life *in vivo*, as the phosphate group (or their mimetics) is essential for recognition and binding to the M6P receptors. Methyldifluoro-phosphonate is also isosteric of the phosphate group, and this is a key factor for the binding with the receptor [185]. It is well known that $CH_2PO(OH)_2$ can successfully mimic, the natural

phosphate; in addition, M6Pn can bind CI-M6P as well as M6P [185], so it reasonable to aspect that the difluoromethyl phosphonate will work as well as the phosphate. The spatial parameters are similar to those of the phosphate; moreover, the presence of two electron-withdrowal fluorine atoms will increase the polarity, increasing the acidity of the phosphonate group (second pKa 5.4, Phosphate second pKa 6.4). Hence, it is right to presume that the binding affinity with the receptor can benefit from the increased activity because it will increase the ionic interaction ligand-receptor.

The presence of two fluorine atoms per molecule of Mannose is very interesting because this could allow eventual MRI application[192] and tumour detection; in fact, as CI-M6P is over-expressed in most of the solid tumours [175, 188], this could represent a novel probe for a new imaging target.

The anomeric position of mannose will be alfa-glycosilated with propargyl alcohol, (Compound **82** Scheme 17) and this will permit to bind it to appositely modified proteins by Huisgen cycloaddition.

The second idea involves the study of the multivalent receptor binding; it is known that monophosphate displays 10-fold lower M6P/CI-M6PR binding affinity than the bis (phosphate) ligand. However, the previous studies to find the correct interphopshate distance are based on compounds where the two M6P are linked together through an aliphatic chain[182] that is very flexible, so the real distance between the mannose groups could be different from

the one calculated by molecular modeling; in fact, the different affinity values obtained using different chain lengths are not very far from the one obtained with M6P alone.

The revolutionary idea is to use a protein as molecular ruler: Np276 is a rigid protein with a distance of 5 Å between every rings of the "spring" [193]. By mutating Np276 with azide-(Aha), in specific positions it will be possible to tag the protein with two M6P analogues spaced from around 20, 25 and 30 Å (Fig. 26). Also single mutation will also be made: in this way it will be possible to check the divalent binding, and we expect the highest affinity when applying the correct interphosphate distance, with the largest affinity difference between the mono- and the di-glycosilated protein.

Using the same protein scaffold, the only variable will be the distance: if this procedure works, not only the important goal of demonstrating the divalent binding in CI-M6PR will be reached, but also the innovative use of a protein as molecular ruler will be shown.

Proteins tagged with the new mimetic (M6FPn) will also be synthesized.

This will help to understand if the difluoromethyl-phosphonate can mimic well the phosphate group during the M6P/CI-M6PR recognition process. If the answer is yes, this will be a new and interesting M6P analogue.

Moreover, thanks to the peculiar properties of Np276, this new glycoprotein will throw light on the receptor binding mechanism.

As discussed previously, M6P analogues have very important applications, both actual and future. The glycoproteins synthesized would be tested to check if they could selectively target HSCs. As previously described, HSCs were tagged with albumine modified with M6P, without paying attention to the inter-mannose distances; if the divalent binding is really important for the recognition process, we expect to see an improvement with the correct inter-phosphate distance.

The M6P- and M6FP-Np276 will be coupled at the N-term with a fluorescent probe (such as Alexa Fluor). This will help during the biological tests, making it easy to check the uptake and cell targeting.

What I aspect from these tests is that if the divalent binding is reached, the largest difference will be detected between the mono- and the right di-glycosilated protein, while maybe a smaller difference will be detected between the three different distances.

In the same way it could be possible to also test if they can selectively target lysosomes and tumour cells, and inspire future works as specific drug carriers or imaging agents.

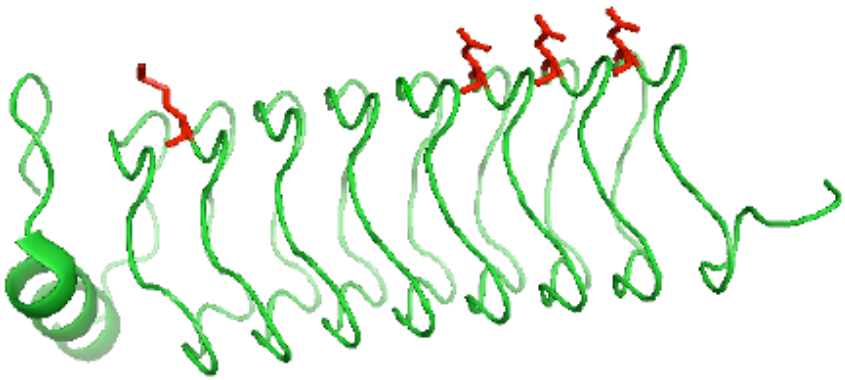


Fig. 26. Np276; in red are evidenced in redthe aminoacids that were mutated into N_3-Ala.

Synthesis of Propargyl-M6FP

The three step procedure reported in the literature[194] for the synthesis of 2,3,4-tri-O-benzyl-D-mannopyranoside **83** (Scheme 17) failed for me; so I synthesized compound **84** protecting Man in position 6 with TBDPS in pyridine, followed by benzylation (**85**) and selective deprotection of the silyl ether with TBAF in THF. Since (**83**). Because the reaction with NaH is not clean, the benzylation is done using benzyl bromide and NaOH as base, and to reach desired yields an efficient stirring is important. This new three step procedure gave better yields than the one reported also in 1g scale obtaining 67% overall yield.

The benzylated intermediate **83** allows modification in position 6. One of the most used methods to insert a difluoromethylenphosphonate group is the replacement of a triflate

group operated by $LiCF_2PO(OEt)_2$ in THF [138]. Unfortunately, this reaction on compound **83** has been very problematic.

First, I synthesized the triflate **86**, subjecting **83** to a triflic anhydride solution in DCM, using DTBMP as base because pyridine is known to substitute OTf [137]. The work-up of this reaction was very important; the triflate **86** was unstable to extraction with DCM and water or $NaHCO_3$ solution, and the best work-up procedure that I found was a fast silica filtration of the reaction mixture, without previous extraction. The evaporation of the solvent in rotavapor has to be done at low temperature (20°C) and then the syrup is dried under vacuum, in ice bath. This step is important because the presence of EtOAc has to be avoided, but too much time increases the decomposition, easily detectable in the colour change (from light yellow to dark).

After many experiments (Fig. 27) I found that in order to have the best yields (around 60%) for the synthesis of compound **87**, a -78°C THF solution of diethyl difluoromethylphosphonate has to be added by cannula to a cold solution of LDA and HMPA in THF; it is very important that both the solutions are cold because the forming carbo-anion is very unstable. After stirring for 20 minutes at -78°C THF solution of the freshly prepared triflate 4 was **86** is added slowly by cannula, and finally the solution was stirred for 20 minutes before being quenched.

LDA and the other reagents have to be fresh. Dry solvents have been used, and the glasses were dried at 150°C and cooled down

under vacuum/N_2. To increase the yield it is better to add the triflate around 20-25 min later than the addition of the difluorophosphonate to LDA, then wait only 2 min or produce the carbanion in situ; from the other side waiting too long increases the rate of the carbanion's decomposition.

I found $LiCF_2PO(OEt)_2$ very unstable, and it decomposes easily; thus, every addition has to be done slowly, by cannula, between -78°C solutions. 1-2g scale reactions gave better results than those on a smaller scale, probably because the moisture variable became less relevant.

4 [mmol]	Tf_2O [eq]	DTBMP [eq]	S/[ml]	Work-up	A/B	LDA [eq]	DIPA/HMPA/nBuLi [eq]	$CHF_2PO(OEt)_2$ [eq]	T_1 [°C]	T_2 [°C]	t_1 [min]	[%]	*
0,234	1,5	1,55	DCM/3	$NaHCO_3$/ice	A	3,5	/	3,6	/	/	/	/	
0,200	1,05	1,05	Et_2O/3,6	Filtration	B	3,5	/	3,5	rt	rt	10	/	
0,183	1,2	1,2	THF/2	/	A	3	/	3	/	/	/	/	*a*
0,081	1,1	1,25	DCM/0,5	H_2O	A	2	/	2	/	/	/	/	
0,185	1,1	1,1	Et_2O/3	Filtration	B	4	/	4	rt	rt	10	6	
0,107	1,1	1,1	Et_2O/3	Filtration	B	2	2 HMPA	2	rt	rt	11	4	
0,181	1,1	1,1	Et_2O/3	Filtration	B	/	4	4	rt	rt	13	7	
0,093	1,2	1,6	DCM/1,1	Silica	A	/	2	2	/	/	/	10	
0,148	1,2	1,6	DCM/1,2	Silica	B	/	3,5	3,5	rt	-78°C	5 – 10	25	
0,148	1,2	2,6	DCM/1,2	Silica	B	/	4	4	rt	-78°C	2	10	*b*
0,148	1,3	1,6	DCM/1,2	Silica	A	/	5	5	/	/	/	8	*c*
0,119	1,2	1,6	DCM/1	Silica	B	/	3,5	3,5	rt	-78°C	20	12	
0,122	1,5	1,6	DCM/1,2	Silica	B	/	3,5	3,5	rt	-78°C	10	14	
0,378	1,5	1,7	DCM/3	Silica	B	/	3,5 + 1,5	3,5 + 1,5	rt	-78°C	30	15	*d*
0,371	1,5	1,7	DCM/3	Silica	B	3,5	3,5 HMPA	3,5	-78°C	-78°C	2	25	*e*
0,404	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	40	
0,487	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	35	13	
0,426	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	33	12	
0,404	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	20	30	
0,404	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	30	
0,835	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	34	
1,332	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	50	
1,852	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	62	
3,354	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	60	

A. *$LiCF_2PO(OEt)_2$ in situ production.* **B.** *ROTf added to $LiCF_2PO(OEt)_2$ solution after a time t_1.*
% yield is calculated over two steps
a. I used THF to avoid any work-up, but cooing the solution at -78°C gave a kind of "gelification"-precipitation.no stirring possible
b. Mistake with the cannula addition of the triflate solution: it was not done drop by drop, but fast.
c. Slower elution starting from 9:1 to 7:3 increased the lost of triflate compound in column, and this compromised the yield. 1h reaction didn't change.
d. LDA and difluoro-phosphonate are added in two steps.
e. From now i'm using a new fresh bottle of LDA from Sigma.

Fig. 27. Summary of the most important experiments

Once I found the conditions for the synthesis of compound **87**, I went further looking for the right deprotection conditions. My early idea was to cleave benzyl and ethyl groups, one pot using TMSI, and then do a Fisher glycosilation on the anomeric position with propargyl alcohol, but the deprotection reaction with TMSI did not work.

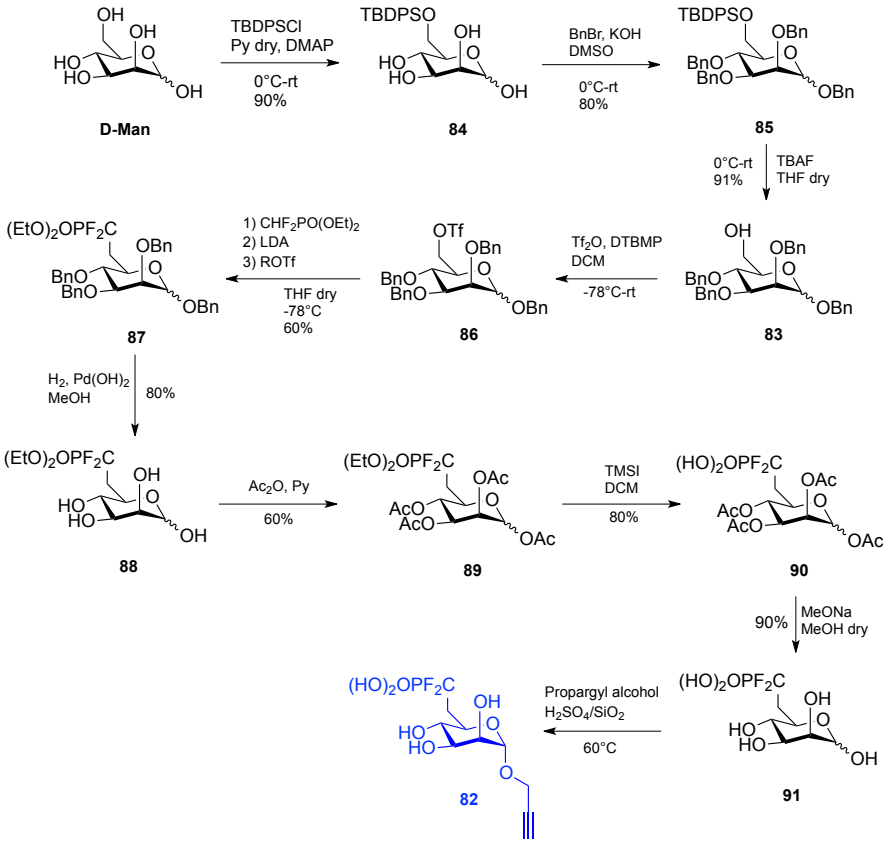
Glycosyl iodides can be synthesized treating a peracetilated carbohydrate with TMSI[195], so I thought to cleave the benzyl groups from compound **87** by hydrogenation, then acetylate the hydroxyl groups and deprotect the phosphonate with TMSI; in these conditions I should also have the glycosyl iodide, which would be suitable for the propargyl glycosilation.

Deprotection of benzyl groups was performed in EtOH with $H_2/Pd(OH)_2$ affording **88** in quantitative yields. Acetylation using I_2 as catalyser is described as the fastest acetylation method[196], but when applied on compound **88** was necessary to add much more iodine than the catalytic amount, so I performed the acetylation in Py/Ac_2O with DMAP as catalyser, giving **89**. The next reaction with TMSI worked cleanly, giving only the deprotect phosphonate, but the compound was still acetylated on the anomeric position, even adding 20 eq of TMSI over 4 hours!! Quenching it with MeOH and then pH neutralization adding NH_4OH solution gave compound **90**. The BF_3OEt catalyzed glycosilation was not successful. Therefore, I chose to deprotect the acetyl groups and glycosilate it by H_2SO_4-silica catalyzed Fisher glycosilation.

The reactivity of this compound is strongly influenced by the presence of the difluorophosphonate; probably the anomeric position is deactivated by the difluorophosphonate and this could explain why the glycosilation did not work and why the TMSI (even if used in large excess) did not deacetylate the anomeric position, while it is known that in peracetylated Man the anomeric position is iodinated in these reaction conditions. For the deacetylation 3 eq of MeONa were necessary, otherwise the reaction was too slow. The resulting product **91** is no longer soluble in MeOH, so it is present as a white precipitate that can be collected by centrifugation. The surnatant is evaporated and the solid suspended in MeOH; after centrifuge it is possible to collect more product, then this procedure is repeated, but the solid obtained is step by step less pure.

Compound **91** is suspended in propargyl alcohol and stirred at 60°C for 10 min, then under stirring H_2SO_4-SiO_2 is added; for the same reaction with mannose 60mg/mmol are used, while in this case 250mg/mmol were required, probably because of the phosphonate. The reaction is done twice: the first time it is filtered and purified by a flash cromatography, and then separated from silica using an ion exchange resin; while the second time the reaction is filtered immediately on ion exchange resin, but I obtained only a few mg and not clean. The main problem is the purification step, because following the reaction by MS it is possible to see the starting peak disappear and the formation of only the product **82** signal (m/z (EI-): 331.0392 [M-H]), Probably the best

way is to work it up with a very short silica filtration and then purify it by HPLC.



Scheme 17. Synthesis of α–propargyl D-mannose-6-difluoromethylphosphonate.

Synthesis of Propargyl-M6P

Starting from D-Mannose I synthesized compound 92 (Scheme 18) by Fisher glycosylation using H_2SO_4-SiO_2 (60mg/mmol) as catalyzer, in propargyl alcohol[197]. Temperature and reaction time were important variables, and running the reaction at 65°C for 3

hours and with freshly prepared silica, gave the desired compound with 74% yield.

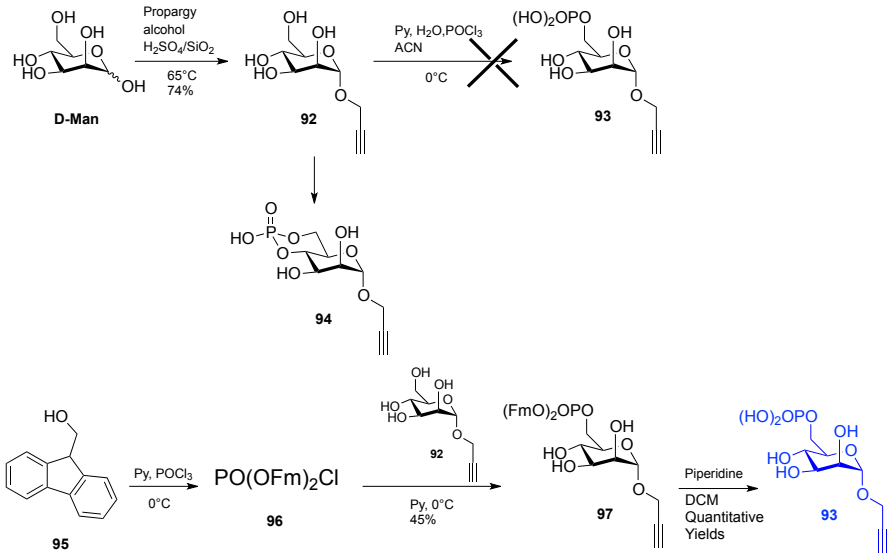
Starting from compound **92** I tried to synthesize **93**, modifying a procedure that uses phosphoric acid, BuN_3 as base and DMAP as catalyser[198]. This procedure is used for the phosphorylation of nucleosides, but it requires an azeotropic solvent and has to be done in reflux to remove water. Since I was using a deprotected sugar, I tried to do it with a larger amount of catalyzer, in the best solvent that the authors found (dry DMF), but without reflux, to achieve a selectivity for the primary alcohol; to remove water I used molecular sieves, but after 2 days, even using sonication, I could not observe product formation.

Next I tried to synthesize **93** following the procedure described in an article for a mild phosphorylation in position 6 of deprotected mannose (64% reported) using $POCl_3$[199]. The reaction was working very clean, giving only one spot tlc detectable, but the obtained product was not the desired one, since I had probably a phosphate bridge between the hydroxyl groups in position 6 and in position 4 (**94**).

Therefore, I tried a reaction of $POCl_3$ with two equivalents of 9-fluorenyl-methanol (**95**) and then added directly compound 11 to the reaction mixture α-propargyl-mannose **92**. I chose 9-fluorenyl-methanol as protective group because in theory its deprotection condition will not affect the propargyl stability. In the first experiment I checked the formation of the desired specie $PO(OFm)_2Cl$ (**96**),

following the reaction in TLC and checking the TLC spot by mass spectra.[200] Once I found the correct conditions, I repeated the reaction adding compound **92** to the mixture after 40 minutes. After one hour the reaction was quenched adding H_2O, and it gave compound **97** with 45% yield.

The protective groups were cleaved by piperidine in 5 minutes, affording compound **93**.



Scheme 18. Synthesis of α–propargyl D-mannose-6-phosphate.

According to my bibliography research this is a new phosphorylation method, selective for the position 6 in unprotected carbohydrates, with the advantages of a fast and efficient protective group cleavage that does not require hydrogenation, which will be unsuitable for double and triple bond substituents.

While the mass observed (in H_2O solution) is correct, I noticed something curious performing the NMR analysis. After the piperidine deprotection, the solvent was evaporated at reduced pressure, then dissolved in D_2O and analysed, but the ^1H spectra did not show the signal of the propargyl proton on the triple bond.

To confirm the hypothesis that due to the piperidine the acid proton had been exchanged with deuterium, the D_2O solution was diluted in H_2O and lyophilized the procedure was repeated three times, then the solid was dissolved in deuterated DMSO. The ^1H spectra now showed the missing proton (Fig. 28) and HMBC analysis confirmed the presence of the propargyl group in anomeric position.

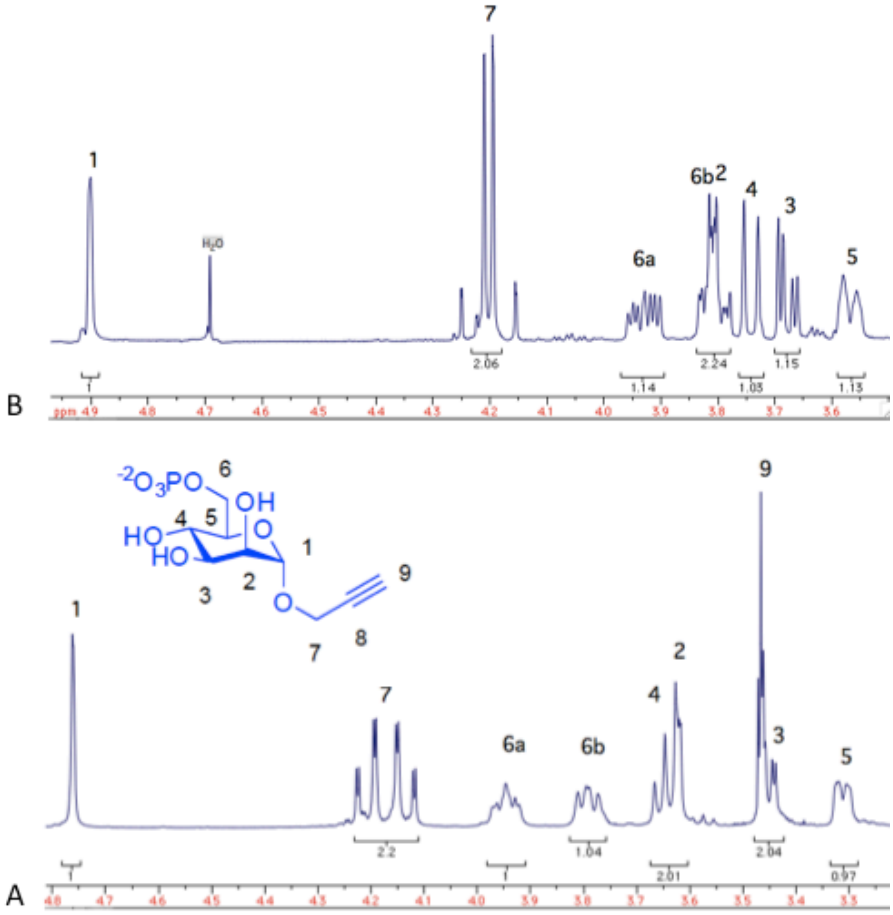
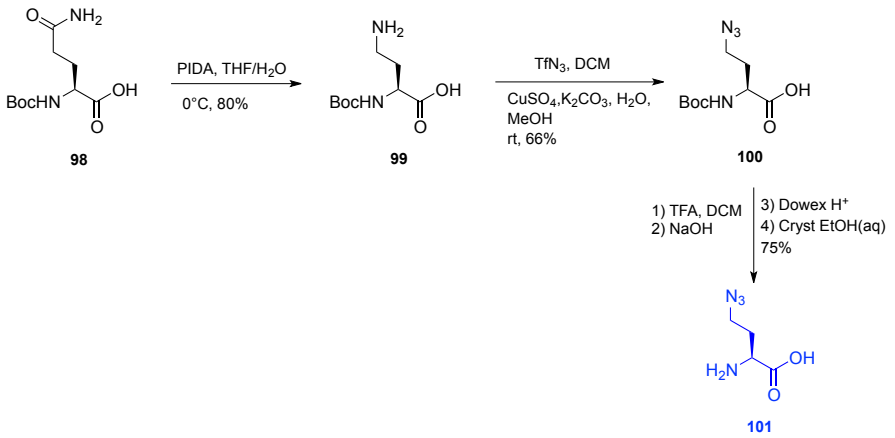


Fig. 28. B) ^{1}H-NMR spectra of compound **93** in D_2O. A) ^{1}H-NMR spectra of compound **93** in DMSO after treatment with H_2O and freeze-drying.

Synthesis of L-Homoazidoalanine

L-Homoazidoalanine **101** (Scheme 18) is synthesized starting from N-tert-Butoxycarbonyl-(S)-glutamine (**98**), using PIDA overnight. The reaction of azido transfer worked well overnight at rt, affording **100**.[201]

The last step is the Boc deprotection in TFA/DCM, followed by ion exchange treatment to remove the triflic acid and crystallization in EtOH to have Haa **101** as a white crystal.



Scheme 19. Synthesis of Homoazidoalanine.

Mutagenesis experiments

I planned to modify Np276 in two positions (Fig. 29) so that the distance between them could reach 30Å and based on the information that we have about Np276, this would be possible only by modifying amino acids 41-161. Modifications too close to the stop codon or to close to the end of the protein were not admitted, because the stop codon would interfere with the modification and in the other way the end of the protein would be too flexible to assume a rigid distance between the two modified sites.

Because the ruler available in our group was already mutated in position 41-141, in order to obtain the desired plasmids, doing the smallest number of modifications, the plan was the following.

Introduce M61I and D41M:

_ For 20Å ruler: from Met 61, 121

_ For 25Å ruler: from Met 61, 141

_ For 30Å ruler: from Met 61, 101. It also needs D161M and M101I.

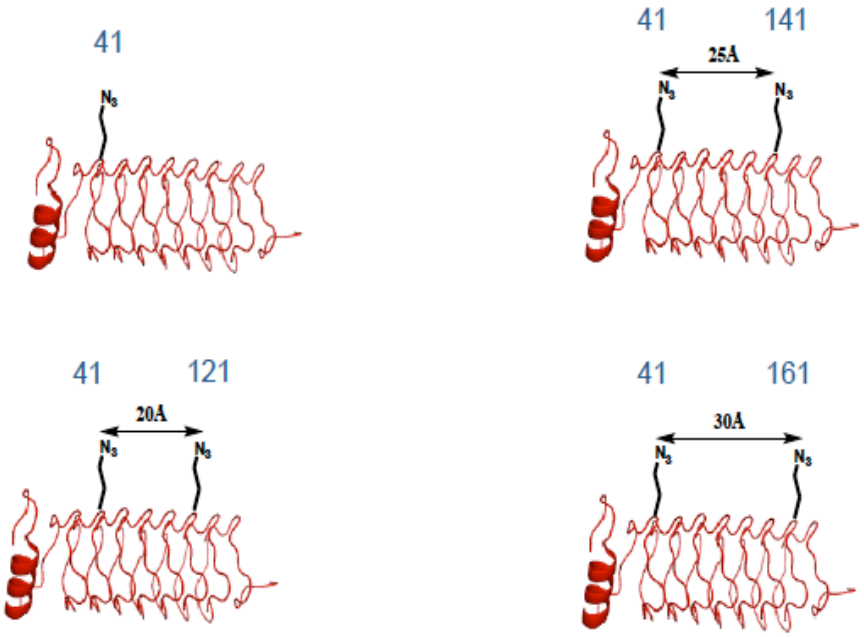


Fig. 29. Np276/molecular ruler: single mutated (control) and 20, 25, 30Å rulers.

The QuikChange® II XL site-directed mutagenesis kit was used to mutate the protein in the specific position. After the PCR cycles, I was running DNA gel to check if the mutagenesis reaction happened: if there was not any mutagenesis, no spot was

detectable on the gel; if the gel was good, I used the obtained solution for the transformation onto XL10 Gold Cells. After the overnight culture, the DNA was extracted using a specific protocol and then submitted to the GeneService Oxford to check and confirm the presence of the desired mutations.

I report here the sequences of the desired plasmids and the primers used for the mutations.

❑ **For 20A ruler: From Met 61, 121**

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V
D L R G A V L E N I N L S G A I L H G A Met L D E A N L Q Q A N L S R A D L S G A
T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A
Met L A A A N L E Q A I L S H A N I R E A D L S E A N L E A A D L S G A D L A I A
D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

D41M
Forward: 5'
CGAGACTTTAGTATCGTTATGTTGAGGGGTGCAGTCTTG 3'
Reverse: 5'
CAAGACTGCACCCCTCAACATAACGATACTAAAGTCTCG 3'
M61I
Forward: 5'-CATGGGGCGATTCTAGATGAAGC-3'
Reverse: 5'-GCTTCATCTAGAATCGCCCCATG-3'

*www.bioinformatics.org/**primerx/***

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V
Met L R G A V L E N I N L S G A I L H G A I L D E A N L Q Q A N L S R A D L S G A
T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A
Met L A A A N L E Q A I L S H A N I R E A D L S E A N L E A A D L S G A D L A I A
D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

NP276 M21/61I K25/94/123A D41M N121M T201K 20 Å ruler

❑ **For 25A ruler: From Met 61, 141**

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V
D L R G A V L E N I N L S G A I L H G A Met L D E A N L Q Q A N L S R A D L S G A
T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A
N L A A A N L E Q A I L S H A N I R E A Met L S E A N L E A A D L S G A D L A I A
D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

D41M
Forward: 5'
CGAGACTTTAGTATCGTTATGTTGAGGGGTGCAGTCTTG 3'
Reverse: 5'
CAAGACTGCACCCCTCAACATAACGATACTAAAGTCTCG 3'
M61I
Forward: 5'-CATGGGGCGATTCTAGATGAAGC-3'
Reverse: 5'-GCTTCATCTAGAATCGCCCCATG-3'

*www.bioinformatics.org/**primerx/***

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V
Met L R G A V L E N I N L S G A I L H G A I L D E A N L Q Q A N L S R A D L S G A
T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A
N L A A A N L E Q A I L S H A N I R E A Met L S E A N L E A A D L S G A D L A I A
D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

NP276 M21/61I K25/94/123A D41M D141M T201K 25Å ruler

❑ **For 30A ruler: From Met 61, 101**

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V
D L R G A V L E N I N L S G A I L H G A Met L D E A N L Q Q A N L S R A D L S G A
T L N G A D L R G A N L S A A D L S D A Met L D N A I L E G A I L D E A V L N Q A
N L A A A N L E Q A I L S H A N I R E A D L S E A N L E A A D L S G A D L A I A
D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

M101I
Forward: 5'-GTGATGCA**ATT**CTTGACAATGC-3'
Reverse: 5'-GCATTGTCAAG**AAT**TGCATCAC-3'
D161M
Forward: 5'
CAGATTTAGCGATCGCGATGTTGCATCAGGCAA ATCTG 3'
Reverse: 5'
CAGATTTGCCTGATGCAACATCGCGATCGCTAA ATCTG 3'

D41M
Forward: 5'
CGAGACTTTAGTATCGTTATGTTGAGGGGTGCAG TCTTG 3'
Reverse: 5'
CAAGACTGCACCCCTCAACATAACGATACTAAAG TCTCG 3'
M61I
Forward: 5'-CATGGGGCGATTCTAGATGAAGC-3'
Reverse: 5'-GCTTCATCTAGAATCGCCCCATG-3'

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V
Met L R G A V L E N I N L S G A I L H G A I L D E A N L Q Q A N L S R A D L S G A
T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A
N L A A A N L E Q A I L S H A N I R E A D L S E A N L E A A D L S G A D L A I A
Met L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

NP276 M21/61I K25/94/123A D41M D161M T201K 30 Å ruler

❑**For Single Met: From Met 61, 101**

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V
D L R G A V L E N I N L S G A I L H G A Met L D E A N L Q Q A N L S R A D L S G A
T L N G A D L R G A N L S A A D L S D A Met L D N A I L E G A I L D E A V L N Q A
N L A A A N L E Q A I L S H A N I R E A D L S E A N L E A A D L S G A D L A I A
D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

M101I
Forward: 5'-
GTGATGCA**ATT**CTTGACAATGC-3'
Reverse: 5'-
GCATTGTCAAG**AAT**TGCATCAC-3'

D41M
Forward: 5'
CGAGACTTTAGTATCGTTATGTTGAGGGGTGCAGTCTTG 3'
Reverse: 5'
CAAGACTGCACCCCTCAACATAACGATACTAAAGTCTCG 3'
M61I
Forward: 5'-CATGGGGCGATTCTAGATGAAGC-3'
Reverse: 5'-GCTTCATCTAGAATCGCCCCATG-3'

*www.bioinformatics.org/**primerx/***

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V
Met L R G A V L E N I N L S G A I L H G A I L D E A N L Q Q A N L S R A D L S G A
T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A
N L A A A N L E Q A I L S H A N I R E A D L S E A N L E A A D L S G A D L A I A
D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

NP276 M21/61I K25/94/123A D41M T201K Met 41

T

❑ **For 25A ruler: From Met 61, 141**

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V
D L R G A V L E N I N L S G A I L H G A Met L D E A N L Q Q A N L S R A D L S G A
T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A
N L A A A N L E Q A I L S H A N I R E A Met L S E A N L E A A D L S G A D L A I A
D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

D41M
Forward: 5'
CGAGACTTTAGTATCGTTATGTTGAGGGGTGCAGTCTTG 3'
Reverse: 5'
CAAGACTGCACCCCTCAACATAACGATACTAAAGTCTCG 3'
M61I
Forward: 5'-CATGGGGCGATTCTAGATGAAGC-3'
Reverse: 5'-GCTTCATCTAGAATCGCCCCATG-3'

*www.bioinformatics.org/**primerx/***

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V
Met L R G A V L E N I N L S G A I L H G A I L D E A N L Q Q A N L S R A D L S G A
T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A
N L A A A N L E Q A I L S H A N I R E A Met L S E A N L E A A D L S G A D L A I A
D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

NP276 M21/61I K25/94/123A D41M D141M T201K 25Å ruler

Fig. 30. Sequences of the desired plasmids and primers used for the mutations.

Also the following “intermediate” plasmids were synthesized:

_ Np276 M21/61I K25/94/123A D121M T201K

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V D L R G A V L E N I N L S G A I L H G A I L D E A N L Q Q A N L S R A D L S G A T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A Met L A A A N L E Q A I L S H A N I R E A D L S E A N L E A A D L S G A D L A I A D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

_ Np276 M21/61I K25/94/123A D41M D61M D121M T201K

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V Met L R G A V L E N I N L S G A I L H G A Met L D E A N L Q Q A N L S R A D L S G A T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A Met L A A A N L E Q A I L S H A N I R E A D L S E A N L E A A D L S G A D L A I A D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

_ Np276 M21/61I K25/94/123A D41M D61M D141M T201K

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V Met L R G A V L E N I N L S G A I L H G A Met L D E A N L Q Q A N L S R A D L S G A T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A N L A A A N L E Q A I L S H A N I R E A Met L S E A N L E A A D L S G A D L A I A D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

_ Np276 M21/61I K25/94/123A D141M T201K

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V D L R G A V L E N I N L S G A I L H G A I L D E A N L Q Q A N L S R A D L S G A T L N G A D L R G A N L S A A D L S D A I L D N A I L E G A I L D E A V L N Q A N L A A A N L E Q A I L S H A N I R E A Met L S E A N L E A A D L S G A D L A I A D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

_ Np276 M21/61I K25/94/123A D41M D61M D101M T201K

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V Met L R G A V L E N I N L S G A I L H G A Met L D E A N L Q Q A N L S R A D L S G A T L N G A D L R G A N L S A A D L S D A Met L D N A I L E G A I L D E A V L N Q A N L A A A N L E Q A I L S H A N I R E A D L S E A N L E A A D L S G A D L A I A D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

_ Np276 M21/61I K25/94/123A D41M D101M T201K (15Å Ruler)

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V Met L R G A V L E N I N L S G A I L H G A I L D E A N L Q Q A N L S R A D L S G A T L N G A D L R G A N L S A A D L S D A Met L D N A I L E G A I L D E A V L N Q A N L A A A N L E Q A I L S H A N I R E A D L S E A N L E A A D L S G A D L A I A D L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N N N L A K

_ Np276 M21/61I K25/94/123A D41M D101M D161M T201K

Met G S S H H H H H H S S G L V P R G S H I D V E A L R Q L Y A A G E R D F S I V Met L R G A V L E N I N L S G A I L H G A I L D E A N L Q Q A N L S R A D L S G A T L N G A D L R G A N L S A A D L S D A Met L D N A I L E G A I L D E A V L N Q A N L A A A N L E Q A I L S H A N I R E A D L S E A N L E A A D L S G A D L A I A Met L H Q A N L H Q A A L E R A N L T G A N L E D A N L E G T I L E G G N

Proteins over-expression

Once the correct DNA sequences were obtained, I performed the over-expression of the proteins. B834D(E3)pLys ultra competent cells were tested, but unsuccessfully.

On the other hand, with B834D(E3) ultra competent cells the over-expression of the Np276 rulers worked successfully. Many colonies were formed on the plates, and choosing one for each plate, after o.n. culture in LB media, it was possible to do a 1L o.n. culture and express Np276 with Haa instead of Met in the mutated positions.

Cells were centrifuged, and after lysis the lysate was purified. First, I purified the proteins (single ruler and 20Å) manually, using a Ni resin to take advantage of the His tag, but I did not have any clean fraction of the protein.

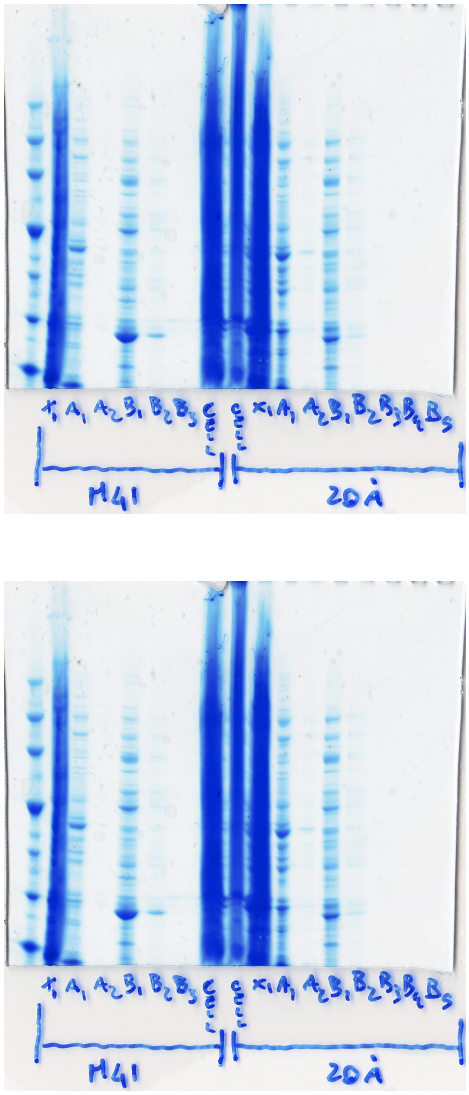
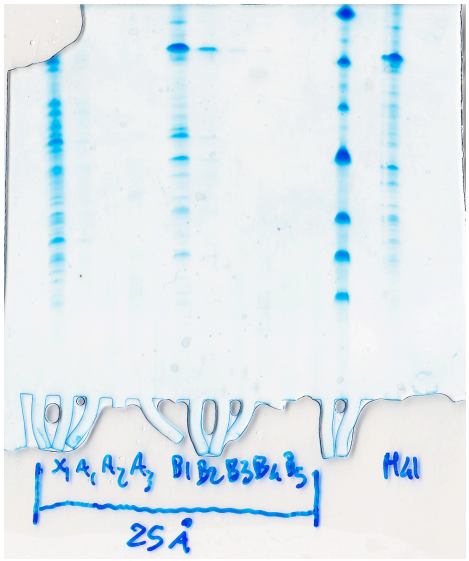
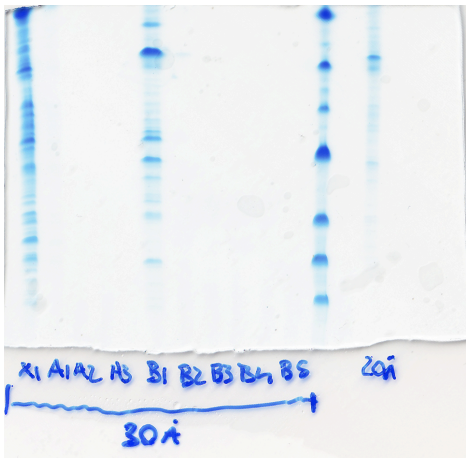


Fig. 31. Gel of the manual purification of Np276 20Å ruler.

I collected the proteins fractions, and I purified the 25Å and 30Å rulers gravimetrically, always using Ni resins; I could remove some impurity, but I did not have any clean fractions.





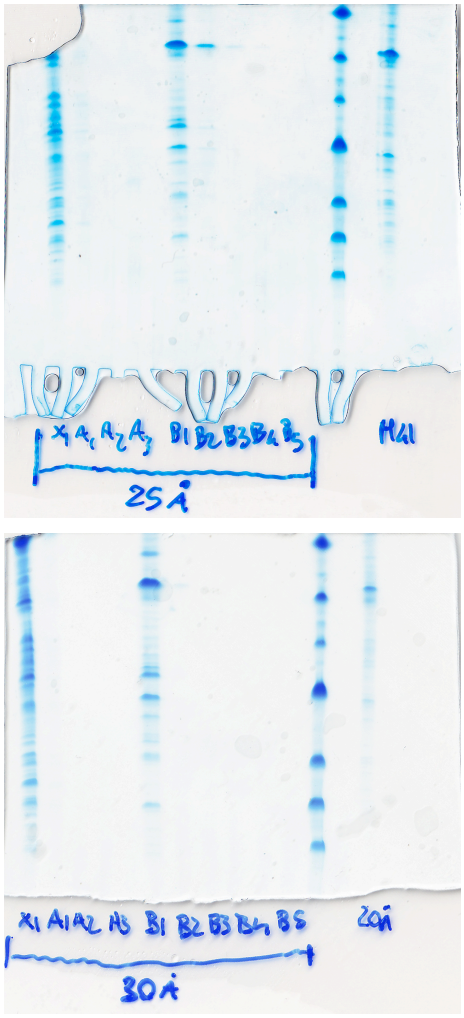
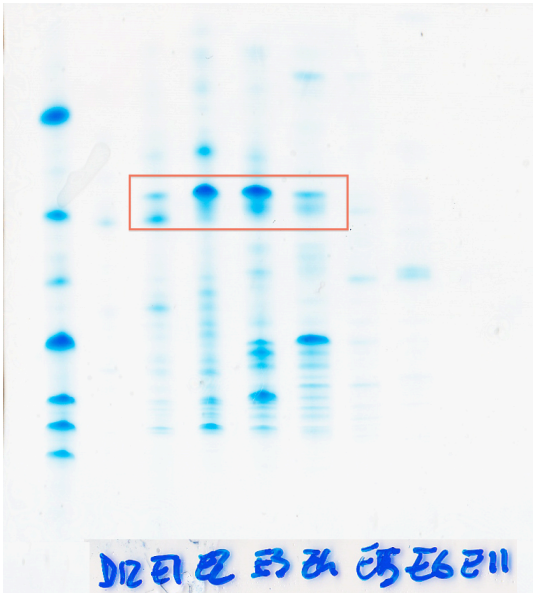
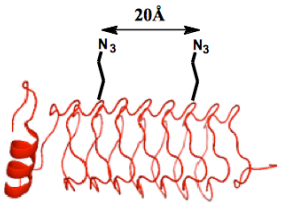


Fig. 32. Gel of the manual purification of Np276 25Å and 30Å ruler.

Thus I collected the fractions containing proteins, and I purified them using an ion exchange resin, and by FPLC it was possible to achieve a nice separation.
Only the 20Å ruler needed a further purification, but using a different elution gradient a better purification was reached.





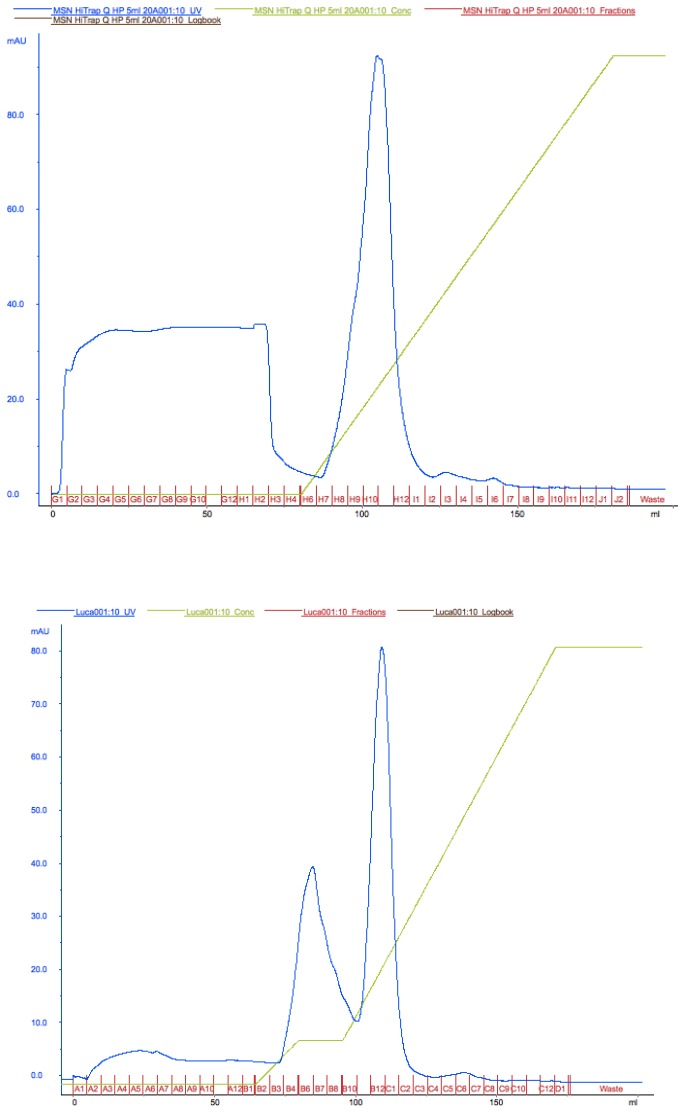


Fig. 33. Gel and chromatogram of the FLPC purification of Np276 20Å ruler.

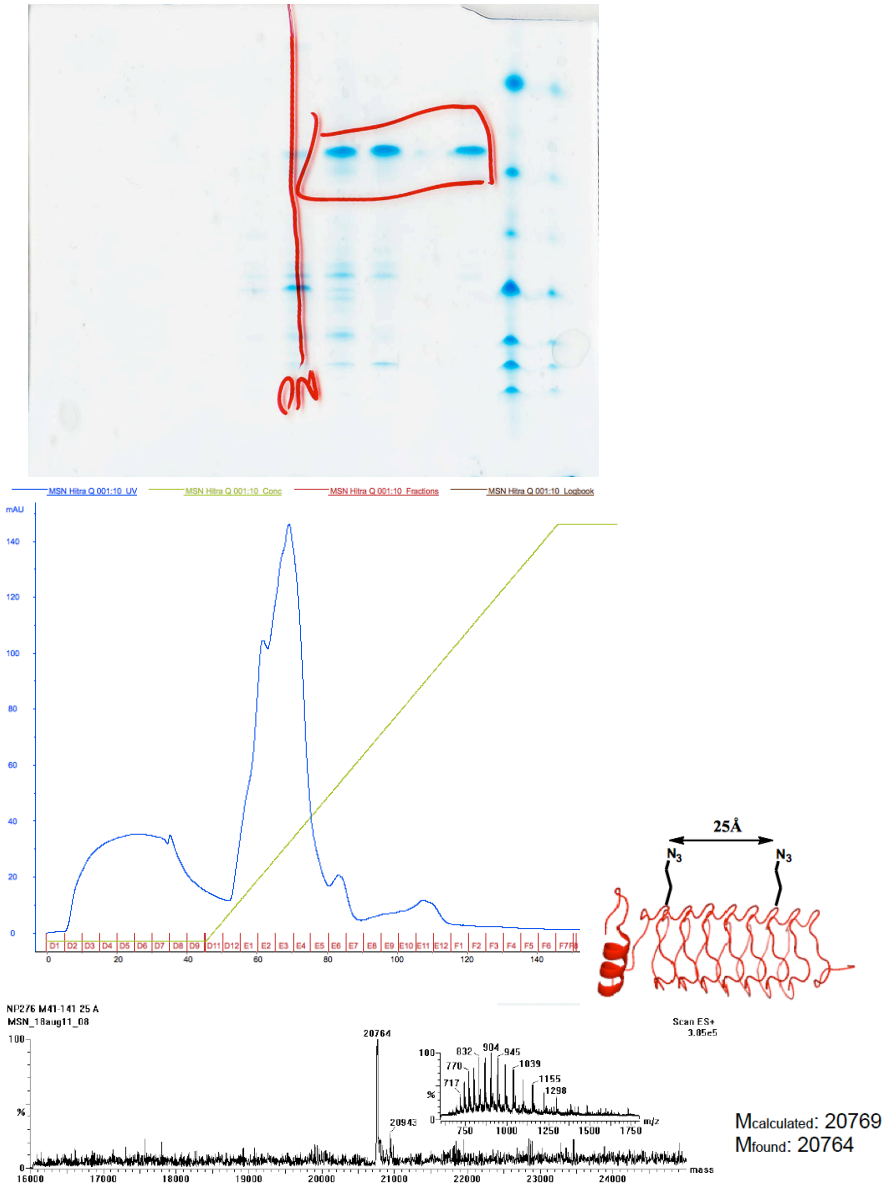


Fig. 34. Gel and chromatogram of the FPLC purification of Np276 25Å ruler and its mass analysis.

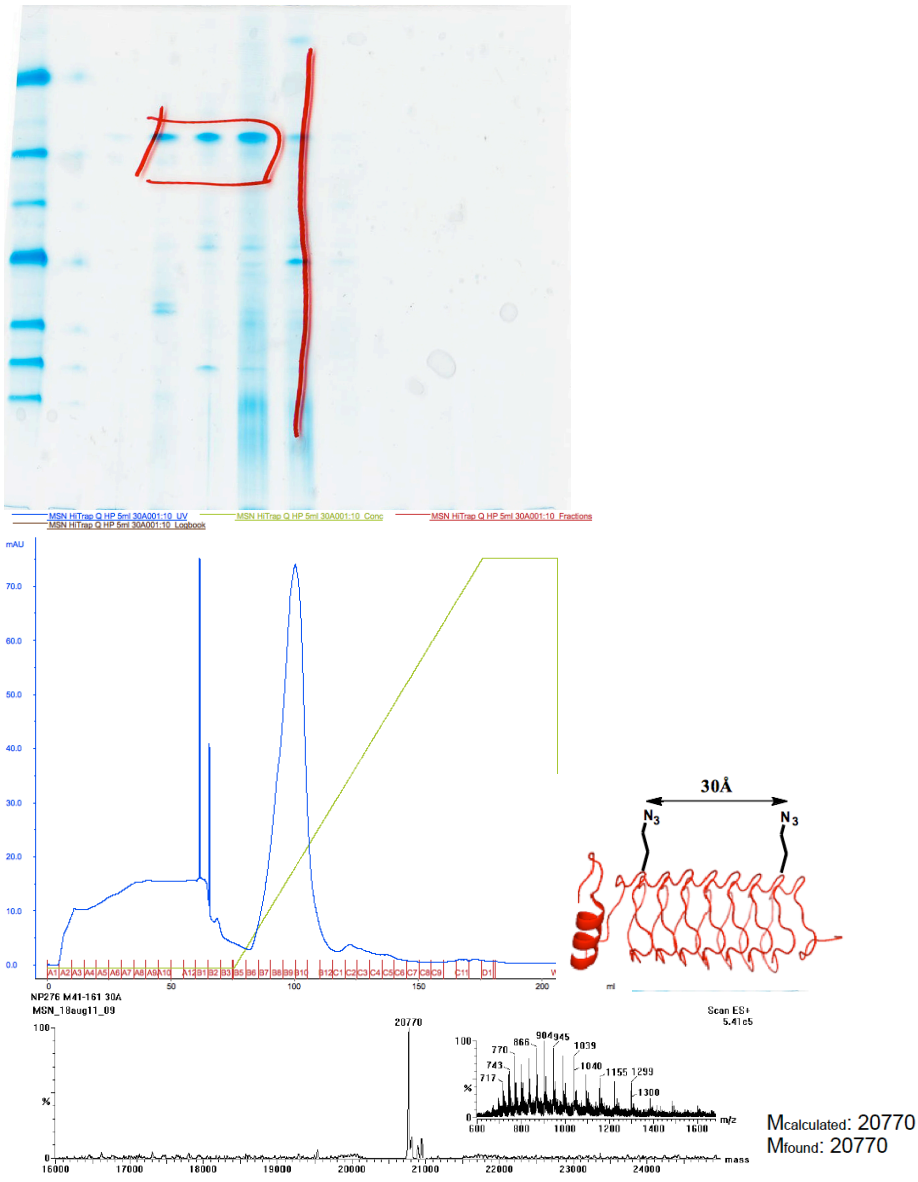


Fig. 35. Gel and chromatogram of the FPLC purification of Np276 30Å ruler and its mass analysis.

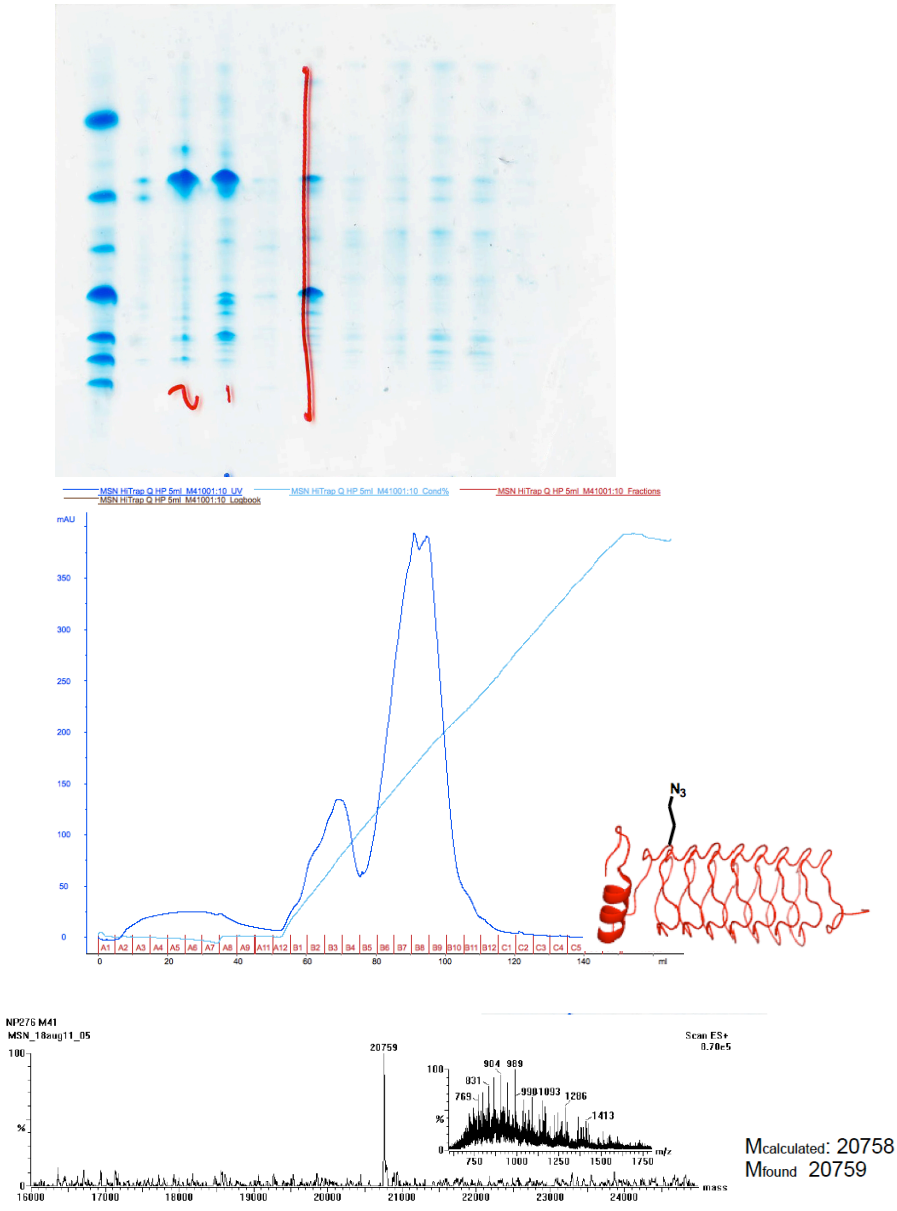


Fig. 36. Gel and chromatogram of the FPLC purification of Np276 single ruler and its mass analysis.

LCT-MS of the fractions confirmed the presence of right protein masses. Only the 20Å ruler gave some problems; after the second purification, I left the sample in the cold room; when I ran the gel again, the protein was no longer present, but it was degraded, as confirmed by the LCT-MS.

Glycosylation reaction on Np276

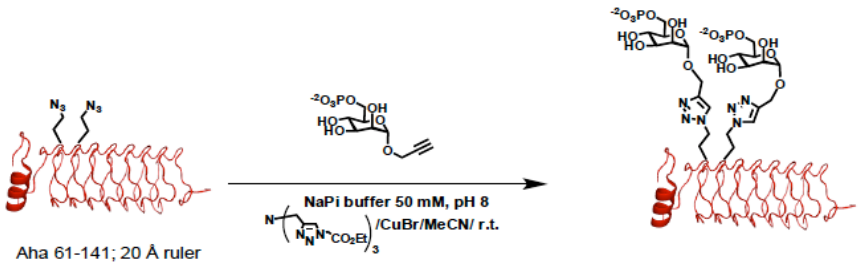


Fig. 37. Reaction of Np276 Aha 61-141 (20Å ruler) with α-propargyl-mannose-6-phosphate **93**

A test reaction of Propargyl-M6P **93** and Np276 Aha 61-141 (20Å ruler) was done to check the conditions of the reaction and to check if my compound was working.

In the first coupling experiments between the propargyl-mannnose-6-phosphate and Np276 20Å, the coupling was quite efficient, but the mass analysis showed 100% hydrolysis of the phosphate group from the glycoprotein.

I thought that the phosphate instability at these reaction conditions could be due to two main factors: the basic pH of the buffer, or the

copper complex that coordinating the phosphate facilitates its hydrolysis.

In order to find the right conditions, wasting proteins, I set up 3 experiments repeating the same conditions of the reaction. In the first one I used only the buffer, to isolate the effect of the pH; in the second one I also added the Cu-complex; in the third one I also added azido-Ala. Samples from the reactions where taken periodically over the 4 hours of reactions and checked by mass.

Unfortunately, the presence of the Cu-ligands was probably quenching the analysis, since no peaks of the carbohydrates were detected; however, for the first experiment (pH 8 buffer) the mass analysis showed no hydrolysis of the phosphate, even after 4h stirring at room temperature. This suggested that the Cu-complex played an important role on the phosphate hydrolysis.

TLC of the samples confirmed that the basic conditions were not enough to hydrolyze the phosphate group, while for the other two experiments a spot of hydrolyzed compound was detected.

Therefore, I repeated the coupling reaction on Np276 (Aha 61-141) 20Å ruler, but increased 10 fold the equivalents of carbohydrate, using propargyl-M6P (**93**). Every 15 min a sample was taken until 1.5 hours of reaction, and the mass analysis showed no hydrolysis of the phosphate and complete conversion already after 15 min, with the best result after 75 min (see experimental part).

6.1. NMR & biomaterials: Introduction

The wide use of the spectroscopy NMR technique performed during my PhD led me to study topics that were not related to carbohydrates themselves, but related to NMR techniques. Therefore, I increased my knowledge and applied NMR spectroscopy in the biomaterial field; in particular I studied by NMR the reactivity of reaction systems widely used for biomaterial production.

Active research lines in our group focus on the functionalization of hybrid biomaterial for bone regeneration. These materials are called "hybrid" because they are constituted by a silica network linked together to an organic moiety. The synthesis of these materials is a multistep one, involving inorganic and organic polymerization; only little information is available on the reactivity of these hybrid systems. Hence, it is evident that knowledge of the system's reactivity will allow material scientists to properly tune the chemical, physical and thus biological property of the material. To the best of our knowledge, it is the first time the formation of a hybrid material (GPTMS/diamino PEG) has been studied in detail by liquid NMR spectroscopy.

Among the numerous silanes used to process hybrid materials, 3-glycidoxypropyltimethoxysilane (GPTMS) is one of the most common ones. It is characterised by a silane moiety, used for the construction of the silica network, and an epoxide ring, useful for organic

functionalization by nucleophilic attack. Due to the high relevance of GPTMS in the synthesis of biomaterial, the next step was the detailed studied of the reactivity of 3-glycidoxypropyltrimethoxysilane as a function of pH, a key variable for the synthesis of hybrid material.

In addition, the GPTMS reactivity with the most used nucleophiles employed for its functionalisation (such as amines and carboxylic acids) have been explored in detail by NMR and mass spectroscopies (data not reported here).

6.2.1. GPTMS/diamino-PEG hybryds: Introduction

Bioactive glass can bond with bone and stimulate new bone growth, making them good candidates for tissue engineering.[202-203] However, they are brittle and to overcome this drawback, inorganic-organic hybrids have been synthesised to improve toughness. Hybrids have molecular scale interpenetrating networks of inorganic and organic components.[204-205] Hybrids are usually synthesised by the sol–gel process, where the organic part is introduced early in the process so that an inorganic (silica) network forms around the polymer molecules.[206] The sol–gel technique has been used to produce several types of hybrid materials for biomedical[207-208] and non biomedical applications.[209] The synthesis is however relatively complex with several chemistry challenges that must be overcome before hybrids will be successful for tissue engineering applications, e.g. incorporation of a suitable polymer and fulfilling the potential of controlled degradation and mechanical properties.

Polyethylene glycol (PEG) is a promising polymer for sol-gel hybrid synthesis as it is biocompatible and biodegradable. PEG/silica hybrids were recently produced by Liu *et al.*,[210] where they added the PEG into the TEOS-based sol at pH 2. No covalent bonds were formed between the components. The rate of loss of PEG from the hybrid was not assessed, but it is likely to be rapid due to the solubility of PEG and the weak interactions between the components.

One way to obtain stronger interaction between the inorganic and organic components in the hybrid material and to gain tight control over the degradation rate and mechanical properties is to obtain covalent coupling between the organic and inorganic components, e.g. through silane coupling agents such as 3-Glycidoxypropyltrimethoxysilane (GPTMS). GPTMS is one of the most common precursors for the preparation of hybrid organic–inorganic materials.[211] The chemistry of GPTMS is quite complex and many aspects are not fully explored and understood.[212] One of the main difficulties in controlling the chemical reactions during the sol-gel process of GPTMS is that several competitive reactions can occur. Usually, when in solution, the epoxy ring of GPTMS (Fig. 38) can be opened by water (affording the corresponding diol), or by nucleophilic attack by –COOH or $–NH_2$ containing molecules,[206] and a bond between the GPTMS and the organic molecule is formed. In the presence of amines as the nucleophiles, GPTMS affords a functionalised trimethoxyorganosilane that can be polymerised with the sol-gel derived inorganic component, e.g. silica species arising from the hydrolysis of tetraethyl orthosilicate (TEOS), through the condensation of Si-OH groups to form –Si–O–S–bonds.[213] The simultaneous polymerisation of the inorganic and organic network is a competitive process;[214] a faster polycondensation rate of the inorganic network hinders the polymerisation of the organic network and *vice versa*.[212] The strict control of the competitive reactions is never simple, since small

changes in the synthesis protocol, e.g. pH, can produce strong differences in the final product.[215] To be able to tune the physical and mechanical properties of the hybrid material it is important to understand the reactivity of GPTMS with the organic component.

Polyfunctional amines are one of the most popular curing agents that are commonly used to fabricate a wide class of thermosetting organic polymers based on the epoxy-amine system. The reactions between epoxy and amino groups to obtain homogenous materials have been the object of extensive research in industry[216]. We proposed the fabrication of novel polyoxyethylene bis(amine)/silica (dPEG/silica) sol-gel hybrids of 35 wt% organic and 65 wt% silica with covalent coupling between the dPEG and silica networks. The primary aim was to produce hybrids with improved mechanical properties for bioactive glasses and bioceramics. Tetraethylorthosilicate (TEOS), GPTMS and dPEG (Fig. 38) were used as components of the reaction system. The bis(3-aminopropyl) polyethylene glycol contains two amino terminal groups available for the reaction with the epoxy group of GPTMS.

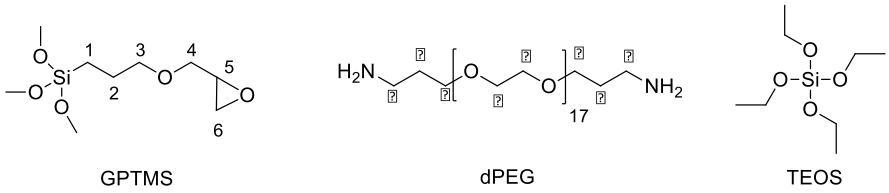


Figure 38. Organic and inorganic reagents for hybrid synthesis: GPTMS = 3-Glycidoxypropyltrimethoxysilane; dPEG = bis(3-aminopropyl) polyethylene glycol; TEOS = Tetraethylorthosilicate.

The hypothesis was that a dPEG/ silica hybrid would have enhanced mechanical properties compared to a bioactive glass and that during the functionalisation of the dPEG the epoxy ring of the GPTMS would open and a covalent bond would form between the $-NH_2$ groups on dPEG and GPTMS (nucleophilic attack). The siloxane pendant groups on dPEG hydrolyse upon immersion in the sol and the newly formed Si-OH groups then condense with the silica network that forms from the condensation of hydrolysed TEOS. The aim was to investigate this hypothesis.

In order to characterise the polymerisation reaction between the organic moieties of the system, that is GPTMS epoxy group and amino groups of the dPEG, a solution-state NMR study was performed; as far as we know, this is the first time that liquid NMR experiments have been performed on such a system.

6.2.1. GPTMS/diamino-PEG hybryds: Results and discussion

Two different ratios of GPTMS and dPEG were combined with a silica precursor to obtain Class II hybrid materials. The interaction between the organic and inorganic components, and the formation of covalent coupling reaction were examined, specifically the nature of the covalent coupling reaction and the formation of the silica network. For coupling to be successful, GPTMS must react both with the polymer and the silica network. If the condensation of the siloxane groups was more rapid than the coupling of the polymer with the GPTMS, then the GPTMS may condense with itself, preventing the formation of a successful hybrid. This would limit the control of the degradation rate and mechanical properties. In order to monitor the condensation reactions, liquid Nuclear Magnetic Resonance (NMR) data was acquired at different time points in the aging process. Aging is the consolidation of a gel in a sealed container post-gelation, prior to the drying step. ^{29}Si and ^{13}C solid state NMR have been widely used in the literature to characterise solid sol-gel materials [217,218,219]. Solution ^{1}H and ^{13}C NMR were used to study the reaction between GPTMS and bis(3-aminopropyl) polyethylene glycol (dPEG) in order to characterise the competition of the organic versus inorganic polymerisation.

The two components (dPEG and GPTMS) were reacted, following a sol-gel protocol, as outlined schematically in Fig. 39. In order to

obtain hybrids with different properties, two different ratios between the GPTMS and the dPEG were used (dPEG:GPTMS 1:2 and 1:4). The sol was prepared by the hydrolysis of the silica precursor tetraethyl orthosilicate (TEOS) in acidic conditions (pH ~ 1). In a separate beaker, a mixture of dPEG-GPTMS was prepared at pH = 6. Afterwards, the TEOS solution was added to the dPEG-GPTMS mixture[206]. The complete sol was poured into Teflon moulds and sealed for the aging at 40 °C for 3 days (Step 4, Fig. 39), after which the tops were unscrewed for a final drying step (Step 5, Fig. 39) at 60 °C for 8 days.

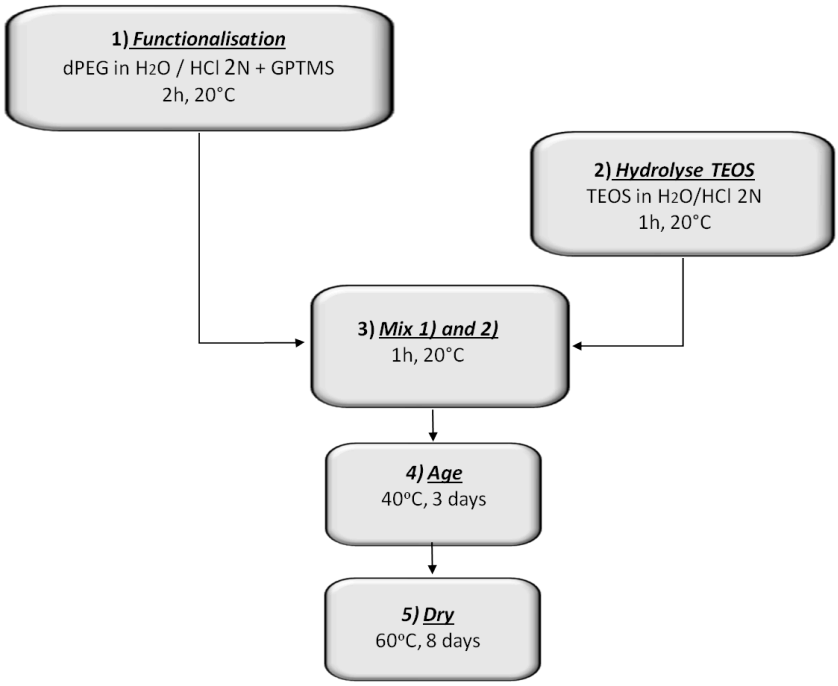


Fig.39. Flow chart showing the reaction and processing steps involved in the synthesis of the dPEG/silica hybrid.

In order to characterise the polymerisation reaction between the organic moieties of the system, that is GPTMS epoxy group and amino groups of the dPEG, a solution-state NMR study was performed; as far as I know, this is the first time that liquid NMR experiments have been performed on such a system. Homo- and heteronuclear 2D NMR experiments of GPTMS (Fig. 40a), and GPTMS in D_2O/DCl at pH 2 (Fig. 40b) and dPEG in D_2O (Fig. 40c), were recorded and used as reference spectra.

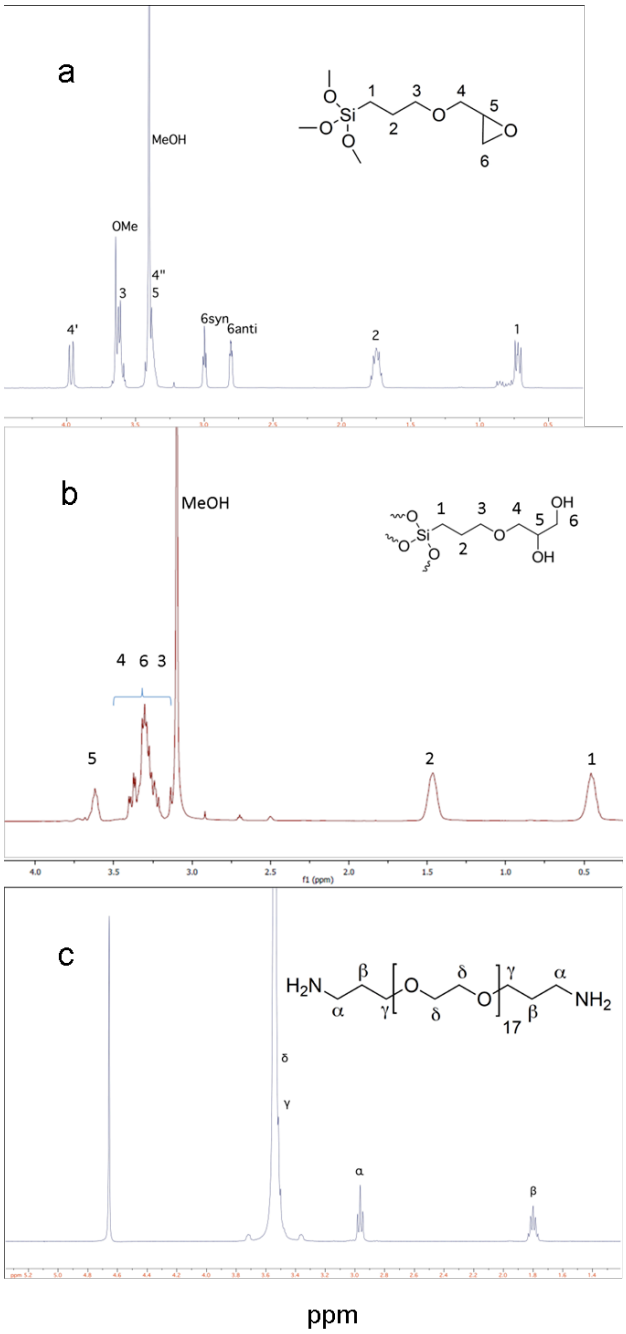


Fig. 40. ^{1}H-NMR spectra and signal assignments of a) GPTMS in D_2O, b) GPTMS D_2O/DCl at pH 2, c) dPEG in D_2O.

The GPTMS-dPEG reaction at pH 6 was monitored by NMR at different reaction times. ^{1}H-NMR spectra of the reactions in deuterated water were recorded at the very end of the 2h functionalisation step (Step 1 in Fig. 39) at room temperature and during the aging step (Step 4 in Fig. 39) at different times (24, 48, and 72 h at 40°C).

Interestingly, for both the 1:2 and 1:4 dPEG:GPTMS mixtures, after 2 h at room temperature, no reaction had occurred at the epoxide ring (H6a δ 2.75, H6b δ 2.94, Fig. 41). In the functionalisation step (Step 1 in Figure 39), the only change detected was partial hydrolysis of the silicon methoxy groups, resulting in the formation of methanol (signal of the MeOH at δ 3.34 ppm) and some Si-OH groups. Thus, the reaction between the epoxide and the dPEG did not occur at this early stage. After 24 h of the aging step at 40 °C (Step 4 in Fig. 39), the NMR data (following the normalisation procedure described in the experimental section) showed that the methoxy groups of the GPTMS were fully hydrolysed (Fig. 41). In particular, the signals due to H2 (1.84–1.60 ppm) and H1 (0.87–0.60 ppm) of GPTMS were much broader and asymmetric, compared to those recorded after 2 h reaction at room temperature. By 24 h, for both reactions (1:2 and 1:4 dPEG/GPTMS molar ratios), the formation of solid material started to be detectable and most of the GPTMS had reacted, forming a silica network. In contrast, almost all the dPEG was still unreacted, as quantified by

NMR (Fig. 41 a-b). After 72 h at 40 °C, the GPTMS signals had almost disappeared, indicating that most of the GPTMS had condensed, and the hybrid had gelled. The solution-phase NMR had characterised the reaction in solution and the final hybrid was analysed by solid-state NMR.

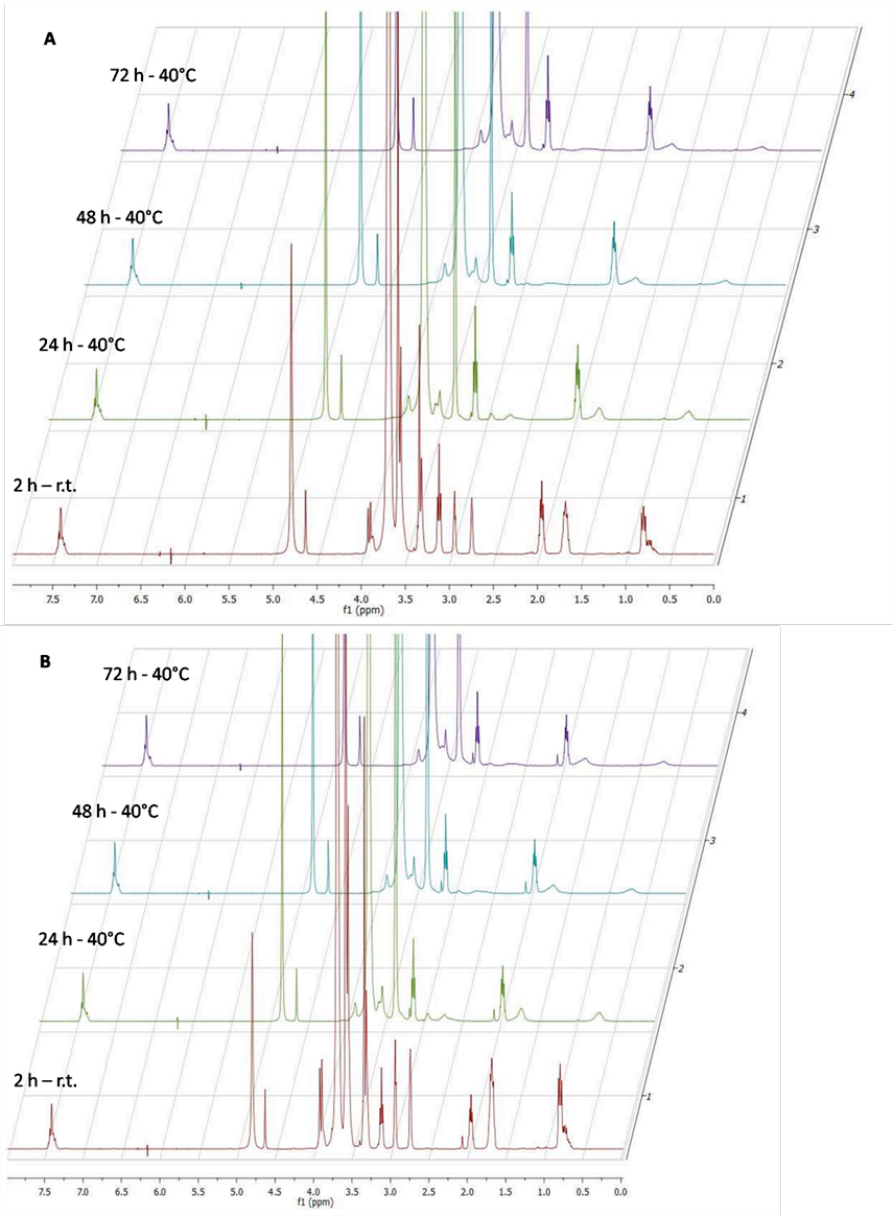


Fig. 41. A: Solution phase ^{1}H NMR of the reaction between dPEG/GPTMS, ratio 1:2; B: Solution phase ^{1}H NMR of the reaction between dPEG/GPTMS, ratio 1:4.

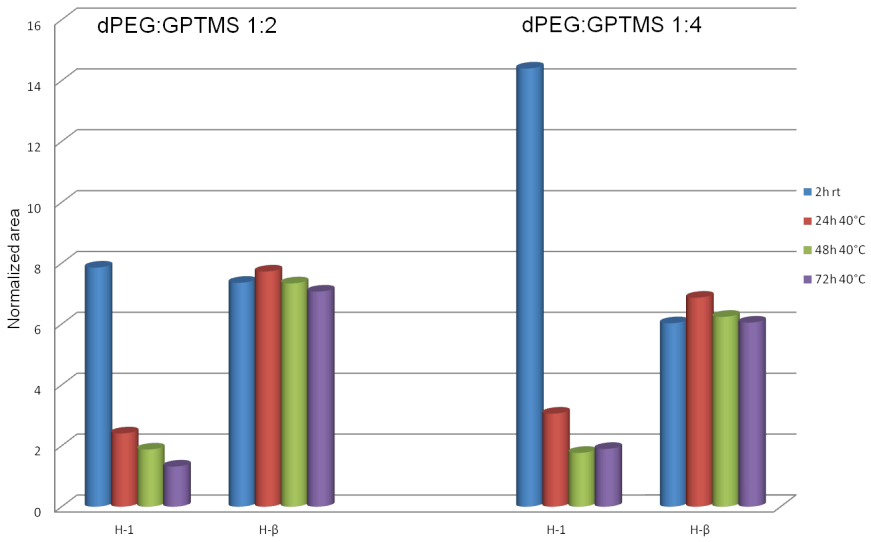


Fig. 42. Amount variations of dPEG and GPTMS in the two reaction systems. For clarity, Arabic numbers are used to designate protons in the GPTMS, while Greek letters are used for dPEG protons.

As previously mentioned, after 24 hours of the aging step, formation of solids occurred, due to some kind of reaction of the species in solution. In order to better characterise the reaction system in solution, using the internal standard (benzyl alcohol) we could detect the reacting species of the system. In fact, the integrals of the peaks of protons of dPEG (Greek letters) and of GPTMS derivatives (Arabic numbers) can be used to quantify the extent of the cross-reaction involving the two species in solution while as a function of time. In particular, H-β (amine functional group) of dPEG (δ 2.01–1.89 ppm) was chosen as characteristic of PEG's peak because it does

not overlap to other signals; H-1 (alkoxysilane group) of GPTMS (δ 0.87–0.67 ppm) was chosen because in that region of the spectra, it is the only signal of all the GPTMS species that is detectable. A decrease of the integrals indicates that the concentration of the compound in solution is decreasing due to the formation of solids; on the contrary, a constant value indicates that the compound does not give cross-reaction with the other reagent, being constantly present in solution, despite the observed solid formation. The variation of the integrals values as a function of reaction time is plotted in Fig. 42: the integral value for the dPEG proton H-β in solution is approximately constant in both reactions with different dPEG/GPTMS ratios (H-β integrals), indicating that the dPEG did not react with the epoxide of GPTMS, at least in the first 72 h. On the contrary, integrals of H-1, decreased as a function of time; moreover, no reaction on the epoxide was detected while the GPTMS was in solution, thus suggesting that GPTMS was aggregating in its epoxide form and any reaction happened afterwards.

In summary, solution-phase NMR has been extremely useful to characterise the reaction features in the first steps of the material fabrication; furthermore, it appears clear that the amine nucleophile is not reacting with the epoxy ring in the solution phase. The final hybrids were produced by drying the aged gels (Step 5 in Fig. 39) at 60 °C for 8 days. The dried transparent crack-free materials were obtained and characterised in terms of morphology and mechanical

properties and by FTIR. FTIR spectra of dPEG/GPTMS hybrids after drying with two different ratios of the organic components (dPEG:GPTMS ratios of 1:2 and 1:4) were collected and compared to a spectrum from a GPTMS/TEOS hybrid (without additional organic components) as a reference. The vibrational modes due to Si–O–Si asymmetric stretching were observed at approximately 1,070 and 1,200 cm^{-1}. Additional peaks were present at 790 cm^{-1} (symmetric Si–O–Si stretching vibration) and at 450 cm^{-1} (Si–O–Si bending modes). Vibration bands from the main silica network content in the hybrid samples between 400 and 1250 cm^{-1} were also present. The additional absorption bands observed around 1070 cm^{-1} were ascribed to the organic moieties of the material, in particular, to the C-O and C–C bonds of the alkyl silane. Other signals between 1350 and 1500 cm^{-1} were due to the C-N stretching vibrations present in plain dPEG and in the hybrid samples, but were absent in the reference material (TEOS/GPTMS). The primary amino groups of the dPEG gave a characteristic broad signal at around 1600 cm^{-1}, which was not present in the spectra of the hybrids, confirming that these groups were involved in covalent bonding with GPTMS.[220-221]

The FTIR data acquired on the solid material, together with the NMR observations in solution, suggest that during the aging time at 40 °C the hydrolysis of methoxy silane groups occurs, with consequent silica network formation. During this step, the epoxide ring did not undergo nucleophilic attack by the amino groups of dPEG.

However, the unambiguous disappearance of the signal of the primary amino groups of the dPEG around 1600 cm^{-1} in the hybrids indicated that the dPEG amino groups were reacted, but most likely the reaction occurred during the drying step at 60 °C.

A new sol-gel hybrid material was synthesised using dPEG as a novel cross-linking molecule. The work was initially based on the hypothesis that reacting dPEG with GPTMS at pH 6 would yield a polymer functionalised with GPTMS molecules, through nucleophilic attack of the epoxy ring by $–NH_2$ groups, and this would enable covalent coupling of the dPEG to the silica network. In order to prove this working hypothesis, for the first time we introduced solution-phase NMR as a relevant tool in order to follow the reaction during material processing. By NMR, we observed, instead, that when the dPEG and GPTMS were mixed together, the siloxane groups on the GPTMS hydrolysed, and the resulting Si-OH groups condensed to form –Si–O–Si– bonds. This means the GTPMS is likely to react with itself and the sol-gel silica network that forms in the TEOS based sol when the dPEG/GPTMS mixture is added to the sol. However, since $–NH_2$ groups were not detected in the hybrids after drying, we can argue that the hypothesised reaction occurred as the water was driven off during drying, producing hybrids with covalent links between the components.

6.3.1. NMR studies of GPTMS reactivity as a function of pH: Introduction

Hybrid organic-inorganic solids prepared by sol-gel processes from organically-functionalised silane precursors represent an important class of engineering materials, since the combination of organic and inorganic moieties affords novel properties and functionalities[206]. 3-glycidoxypropyltimethoxysilane (GPTMS) is one of the most common silane used for hybrid biomaterial synthesis. Hybrid materials synthesised from GPTMS and other precursors such as metal alkoxides and organic crosslinking agents find manifold applications, which include scratch and abrasion resistant coatings on organic polymers,[222] corrosion preventing coatings,[223,224,225] diverse as optical waveguides,[226,227] structured layers for microelectronics,[228,229] and electrically conductive films,[230] adhesive coupling layers for structural joints and fiber-metal laminates,[231] and porous scaffolds (temporary templates) for tissue engineering. During the last decades, several studies have explored the GPTMS-based sol-gel reactions[232,233,234] in order to find any correlation between structure and properties of this hybrid material, allowing a more controlled syntheses of materials with tailor-made properties[235,236]. The polymerisation of the inorganic and organic network is a competitive process;[214] indeed, it has been reported that a faster polycondensation rate of the inorganic network hinders the polymerisation of the organic network, and *vice versa*. Thus there is

the need for a comprehensive characterisation at the chemical level that can reveal the behaviour of the organic and inorganic components, as well as their interplay. In this respect, a comprehensive study of the epoxide-opening *versus* silicon condensation reaction of 3-glycidoxypropyltrimethoxysilane (GPTMS) as a function of pH is still missing, despite initial efforts on this topic.[237] To the best of our knowledge, studies on GPTMS reactivity on its own have been reported at very high pH values,[238,239,240] or at acidic pH in the presence of Lewis acids such as boron trifluoride etherate, titanium chloride,[241] or metal (alk)oxide.[242,243,244]

The NMR and the mass spectrometry study of the GPTMS behaviour as a function of pH was performed in order to compare the reactivity of the silicon condensation against epoxide opening at different pH values. The deep knowledge of the competing reactions (epoxide opening and silicon condensation) that occur in GPTMS might allow the tuning of the silicon network formation and eventually the reactivity against nucleophiles, just by adjusting the reaction pH. This can have profound consequences on the properties of the bulk material. In this context, NMR allows the deep characterisation of the reaction outcome, including hydrolysis and product formation, permitting to address this problem from both the thermodynamic and time-course perspectives. The use of different

NMR-active nuclei also permits to look at the phenomenon form different viewpoints.

6.3.2. NMR studies of GPTMS reactivity as a function of pH: Results and discussion

In this study, I focussed on the behaviour and reactivity of GPTMS in aqueous solution at different pH values, namely 2, 5, 7, 9, and 11. From the technical perspective, a deuterated aqueous solution at the desired pH was prepared using ^{2}HCl 2N or NaO^{2}H 2N, and then GPTMS was added to give a 0.85 M final concentration. Mass spectrometry, ^{1}H, ^{3}C and ^{29}Si nuclear magnetic resonance methods were employed as tools to carefully study the reaction system.

Hetero- and homonuclear 2D experiments of unreacted GPTMS and after complete hydrolysis of the epoxide ring, were used as references for monitoring the reaction time-courses under the different experimental conditions (Fig.43 bottom); The NMR peak assignments were in agreement with the literature data[245].

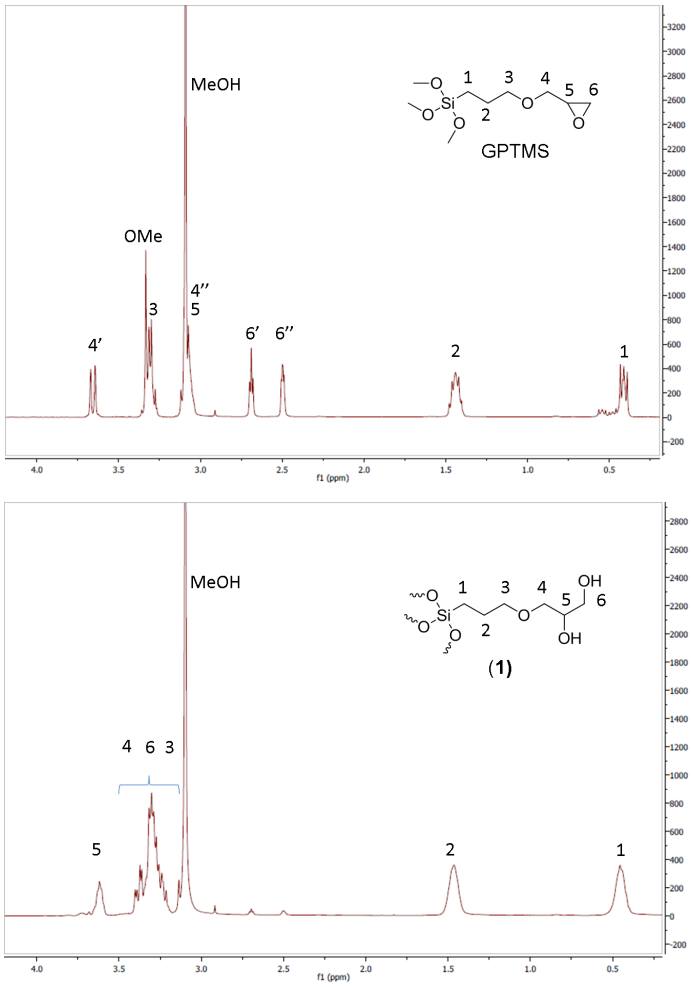


Fig. 43. 1H-NMR of GPTMS (top) and of the fully hydrolysed products (bottom).

From the reactivity perspective, and as an initial step towards the formation of a silicon network by intermolecular condensation reactions, the methoxy groups attached to the silicon atom (3H, s, δ 3.58 ppm) should be hydrolysed to give the corresponding Si-OH moieties. Obviously, this process should be accompanied by the

formation of methanol (3H, s, δ 3.34 ppm). The follow up of the time course of this reaction indicated that the hydrolysis of the alkoxy groups is strictly dependent on the pH of the solution. At acidic pH (2 and 5), already at t_0, both ^{1}H and ^{13}C NMR showed complete hydrolysis of the methoxy groups, with the consequent formation of MeOH. In contrast, at pH 7, according to the ^{1}H NMR data, the methoxy group hydrolysis was much slower: at t_0 no methanol was detected, while after 3 hours, the hydrolysis was 77%. After 28 hours, the methoxy group hydrolysis was complete.

This scenario (pH 2, 5, 7, 9, 11) was also carefully monitored by using a combination of ^{1}H-^{13}C HSQC and ^{1}H-^{29}Si HMBC spectra, along with direct detection of the ^{29}Si signals. Two regions were carefully monitored, the region of the methoxy groups and that of the R-CH_2-CH_2-Si moieties (Fig. 44). After 4-5 hours, only partial hydrolysis products were observed; methanol appeared, with no signs of Si condensation products (only T^0 species). Indeed, no Si-O-Si signals appeared in the ^{29}Si spectrum. The CH_2-Si moiety provided different cross peaks depending on the number of OMe groups attached to the Si atom. After 8 h, the ^{29}Si NMR data indicated that some Si-O-Si signals appeared (T^1 species). After 18 h, T^2 species started to show, but they were still very minor. T^0 species started to decrease after 4 days, while T^1 and T^2 species still remained visible after 14 days. T^3 species emerged after fourteen days.

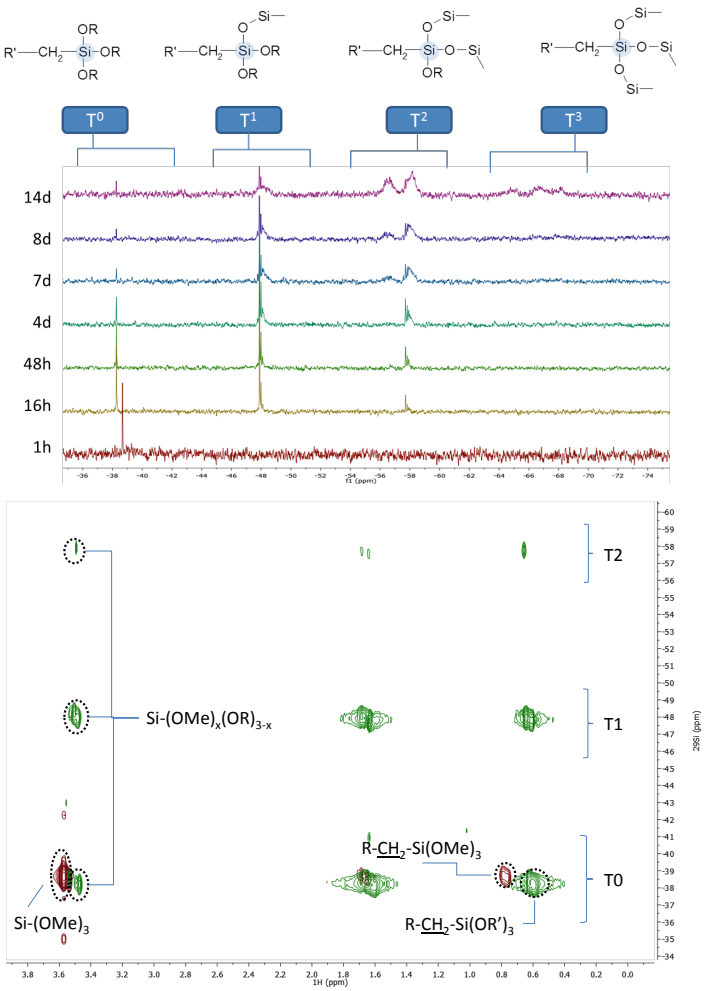


Fig. 44. Top ^{29}Si spectra of the hydrolysis reaction at pH 7. Bottom: HMBC ^{1}H-^{29}Si : red initial, green after 18h

At basic pH (9 and 11), the methoxy group hydrolysis was also slower than at acidic pH. From a macroscopic point of view, at either neutral or acidic conditions, the reaction system remained clear and no formation of solid material could be detected, even after 6 months. In contrast, at basic pH, even after very short times, precipitation of solid entities occurred. The higher the pH, the faster

the precipitation was. The formation of insoluble solid material could be explained in light of the nature of the condensation products that were formed, which were strictly dependent on pH, and could be identified by MS spectroscopy (*vide infra*).

The hydrolysis of the epoxide ring to the corresponding diol is the second and competing reaction that must be taken into account to understand the reactivity of GPTMS. ^{1}H NMR method was employed to easily monitor this process. The hydrolysis of the epoxide caused a significant shift of H6' (dd, J = 4.2, 2.6, ppm δ 2.75) and H6'' (t, J = 4.2, δ 2.94 ppm) to lower fields (3.64-3.46), with concomitant coalescence of the H-6 signals, due to the relief of the conformational constraints of the oxirane ring. Thus, the amount of epoxide hydrolysis can be easily estimated, based on the integrals of the corresponding signals, normalised to the integrals of H2 and H1 signals, which can be semi-quantitatively used as internal reference. Epoxide opening to the corresponding diol was then quantified by using ^{1}H NMR at pH 2, 5, and 7 during the reaction time course. The corresponding plots are shown in Fig. 45. Unfortunately, at pH 9 and 11 the formation of solid materials precluded the proper estimation of the amount of epoxide hydrolysis.

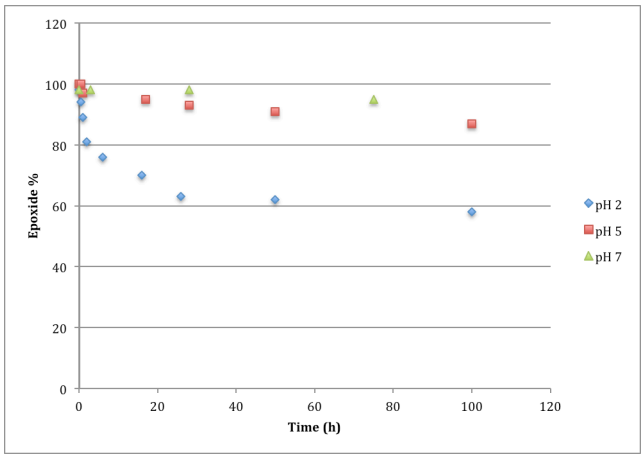


Fig. 45. Epoxide opening as a function of time at pH 2, 5 and 7.

As expected, the rate of epoxide hydrolysis increased as pH decreased from 7 to 2. At neutral pH, this hydrolysis process was very slow, and after 3 days only 5% of the epoxide was converted to the corresponding diol. At pH 5, the presence of acid slightly increased the rate of hydrolysis; after 3 days, 13% of the epoxide ring was hydrolysed. At pH 2 (HCl), the hydrolysis is more accentuated with a strong decrease of the epoxide amount in the first 24 hours: in the first 30 minutes 6 % of the epoxide is hydrolysed, increasing to 11 % after 1 h. After 24 h, the diols accounts for 37 %, and increases to 42% after 3 days. The corresponding NMR tubes at pH 2 were kept for 6 months at room temperature, and after data analysis full hydrolysis of the epoxide moiety to the corresponding diol (Fig. 43, bottom) was observed without formation of any solid material. In addition, it was deduced that the possible competing addition of the chloride ion by

nucleophilic attack during the epoxide opening was not taking place.

At basic pH (9 and 11), the formation of a white precipitate caused strong broadening of the NMR signals. However, the pattern of signals observed in the ^{1}H-NMR spectra, together with the MS data (*vide infra*) provided significant information, showing that the existing compounds were different from those found under neutral or acidic conditions (Fig 46 and 47).

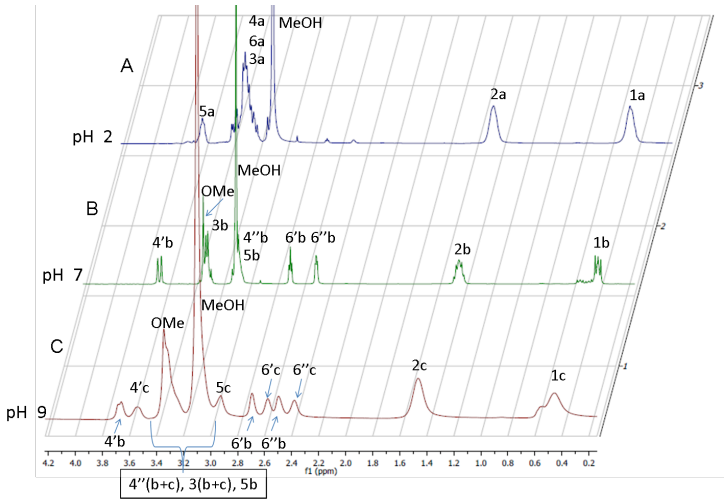


Fig. 46. Comparison of the ^{1}H NMR spectra of GPTMS at different pH values. Subscripts indicate the chemical species that refer to: a: products **1** (see structure in Fig. 43b); b: GPTMS (refer to structure in Fig. 43a); c: products obtained under basic conditions, i.e. **C2**, **D2**, **3A**, **4B**, **4D**, **5B** (see Fig. 52 for structures).

Major differences were observed in the epoxide region in the ^{1}H NMR signals (Fig 46 and 47): the signal assigned to the epoxide

proton at position C-5 (C*H*) upfield shifted from 3.25-3.31 ppm (Fig. 46B, signal 5b) to 3.11-3.21 ppm (Fig. 46C. signal 5c); additionally, these protons at position C-6 (CH_2) shifted from d 2.94 (signal 6'b) and 2.75 ppm (signal 6"b, Fig. 46B), to 2.82 ppm (bs, signal 6'c) and 2.62 ppm (bs, signal 6"c, Fig. 46C). The doublet at 3.91 ppm (signal 4'b, Fig. 46B), which was assigned to one of two protons attached to C-4 was downfield shifted to 3.77 ppm (signal 4'c, Fig. 46C). The ^{1}H-^{13}C heteronuclear HSQC experiment (Fig. 46) also showed that the second proton attached to C-4 (4"b, Fig. 46B) additionally shifted, now overlapping to other signals between 3.24-3.40 ppm (4"c in Fig 46C, and Fig. 47).

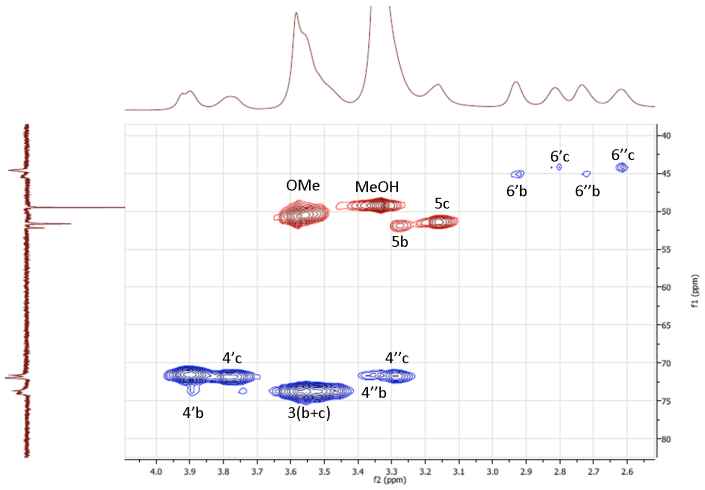


Fig. 47. HSQC experiments at pH9. Assignments are shown.

Interestingly, the two protons attached to C-6 (6'c and 6"c) still displayed two distinct signals (comparing spectra 46B and 46C)

and appeared shifted up-field when the reaction proceeded. This fact contrasts with what was observed at acidic pH, after epoxide hydrolysis, where the two C-6 protons resonated at higher ppm (spectrum 43b) at almost identical chemical shift (Dd around 0.05 ppm). The same trend was observed for the CH_2 protons at position C-4. Additionally, the ^{13}C-NMR spectra of the reaction crude at alkaline pH showed relevant differences when compared to those obtained at acidic pH (Fig. 48).

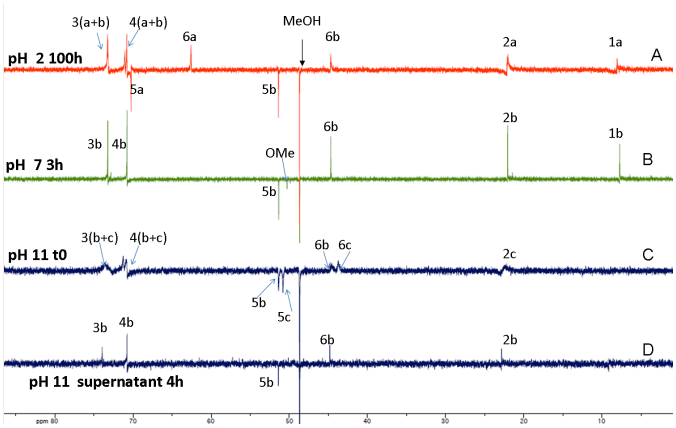


Fig. 48. Comparison of the ^{13}C-NMR spectra of GPTMS hydrolysis at different pH values

At acidic pH (Fig. 48, spectrum A), after 3 h, both the epoxide ring (signals indicated with the b subscript) and the corresponding diols (signals indicated with the a subscript) were simultaneously present. At pH 9 and pH 11 (Fig. 48C), although relatively broad, there were ^{13}C NMR resonance signals symptomatic of the existence of the oxirane ring (signals C5b and C6b). Nevertheless, a new

product is formed (signals 3c, 4c, 5c and 6c). These new signals do not correspond to those observed upon the formation of the diols (see Fig. 48A), and still resonate very closely to those of the starting material (Fig. 48B). In fact (Fig. 48B and 48C), C6b shifted from 44.7 ppm to 43.8 ppm (signal C6c) in basic solution. At the same time, C5 (C5b, Fig. 48B) only shifted from 51.4 to 50.8 ppm (C5c, Fig. 48C). This observation is in sharp contrast to the observations in acidic conditions (Fig. 48A). After hydrolysis of the epoxide (signals C5b and C6b) to the corresponding diol species, the new signals were found at 62.7 ppm (C6a, Fig. 48A), and 70.3 ppm (signal C5a, Fig. 48A). These variations observed in both the ^{1}H NMR (Fig. 46C compared to 46A) and ^{13}C NMR (Fig. 48A and 48C) strongly suggest that under these basic experimental conditions, the formation of the silicate network still involves moieties which maintain the epoxide ring, and that epoxide hydrolysis is not taking place. In fact, the ^{1}H-NMR and ^{13}C NMR (Fig. 48D) spectra recorded on the supernatant displayed fairly sharp signals, which moreover did not show any chemical shift changes in the signals of the epoxide region.

In order to better understand the reactions taking place at different pH as a function of time, mass spectroscopy data were also recorded and integrated with the NMR analysis. This is the first time that silicate network formation of GPTMS is analysed in detail by mass spectroscopy. At acidic pH (pH 2), the mass spectroscopy data strongly supports the NMR-based conclusions, showing the fast

hydrolysis of the methoxy groups to the corresponding silanols. The mass spectroscopy analysis, however, permitted to add one additional feature to the picture, indicating that immediately after methoxy group hydrolysis, silicate network formation (condensation reaction) started. Indeed, after 30 minutes the main peak is due to the dimer A2 (393.3 M+Na), Fig. 52. which displays silanol groups and unreacted epoxide, together with small amounts of trimers C3 (Fig. 52, 569.3 M+Na), tetramers D5 (745.4 M+Na) and pentamers D6 (921.4 M+Na). At this reaction time all the oligomers have the epoxy ring intact, thus indicating that the condensation reaction is faster than epoxy ring hydrolysis. Therefore, at very acidic pH, the reaction starts with the methoxy group hydrolysis to the corresponding silanols, followed by the condensation reaction. Finally, the epoxy ring hydrolysis starts on the formed oligomers, in particular after 3 hours at pH 2 the first signal of epoxide hydrolysis A6 (763.4 M+Na) appears. Generally speaking, at pH 2, epoxide hydrolysis only occurs on oligomerised structures, when the silicon network is already partially formed.

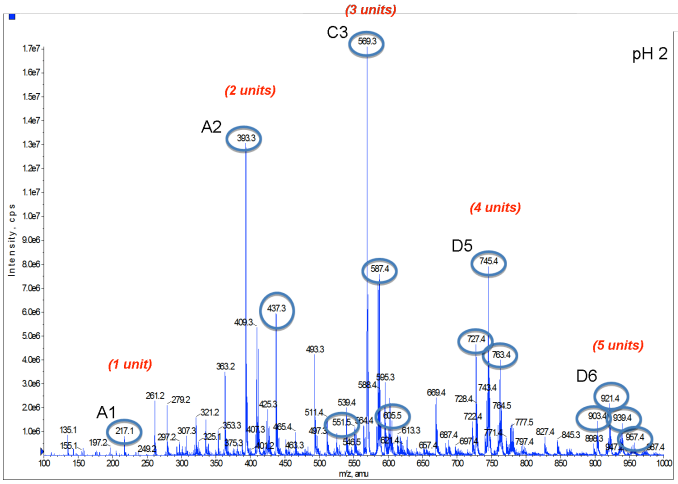


Fig. 49. MS spectra of GPTMS reaction at pH 2 after 24 h.

By increasing the reaction time, the signals due to higher oligomers increased and after 24 hours (Fig. 49) additional peaks were clearly present. The data was analysed in terms of the existence of intramolecular condensation of the oligomers, causing the formation of cyclic oligomers, as compound **C5** in Fig. 52. On both types of oligomers (cyclic or not), the hydrolysis of the epoxy ring took place at slower rate than the condensation reaction. For instance, for tetramer **A6**, after three hours, only one epoxide ring was hydrolysed. Moreover, only traces of pentamer **A7**, with only one diol moiety, were observed. Thus, the combination of MS and NMR data show that both silicate network formation and epoxide hydrolysis took place at pH 2 with different kinetics. Under these experimental conditions, the reaction products were still soluble (with no precipitation) in the reaction medium, and finally no competing HCl addition to the epoxide occurred. At pH 5, the outcome was

very similar to that observed at lower pH (pH 2, silicate network formation and epoxide hydrolysis), but the reactions occurred with reduced speed. At neutral pH, the behaviour of GPTMS was characterised by a slow methoxy hydrolysis. After 30 minutes a part of the clear signals of unreacted starting GPTMS (**D1**, Fig. 52, 259.2 M+Na), additional signals due to the di-methoxylated monomer **C1** (245.2 M+Na), mono-methoxylated **B1** (231.2 M+Na) and non-methoxylated GPTMS **A1** (217.1 M+Na) were visible (Fig. 52 for structures). A part of the clear signals of unreacted starting GPTMS (**D1**, Fig. 10, 259.2 M+Na), after 30 minutes, additional signals due to the di-methoxylated monomer **C1** (245.2 M+Na), mono-methoxylated **B1** (231.2 M+Na) and non-methoxylated GPTMS **A1** (217.1 M+Na) were visible (see Fig. 52 for structures). Therefore, all these compounds contained Si atoms belonging to the structural group T^0 (a silicon atom bonded to a carbon atom and the three remaining Si bonds were non-bridging oxygen bonds). At this pH value, the silicate network formation occurred by condensation of monomers that were not fully hydrolysed to the corresponding silanols, as witnessed by the presence of dimer species in the fully hydrolysed form **A2** (393.3 M+Na), together with mono-methoxylated **B2** (407.3 M+Na), di-methoxylated **C2** (421.3 M+Na), tri-methoxylated **D2** (435.3 M+Na) and tetra- methoxylated **A3** (449.3 M+Na) species (Now, T^1-type Si atoms are appearing). After 2 hours, most of the methoxy groups were hydrolysed and the signal of the trimer **C3** was clearly

evident (569.3 M+Na). After 24 hours (Fig. 50), the demethoxylated monomer **A1** (Fig. 52) was still present, while formation of tetramer **D5** (745.4 M+Na) took place. Both compounds C3 and D5 contain T^1 and T^2 (T^n correspond to an Si atom with a Si-C bond and n bridging oxygen bonds -Si-O-Si-) type Si atoms, thus indicating that these structural groups were increasing, as described above in the ^{29}Si NMR study at this pH. Therefore, at neutral pH, the condensation reaction is predominant, while the epoxide hydrolysis is almost negligible (product **A6**, 763.3 M+Na, is found in the reaction) during the first 24 hours, as also observed by NMR spectroscopy. The condensation towards silicate network formation is much slower than that described at acidic pH. More interestingly, no intramolecular condensation to cyclic dimers and trimers were observed under these experimental conditions.

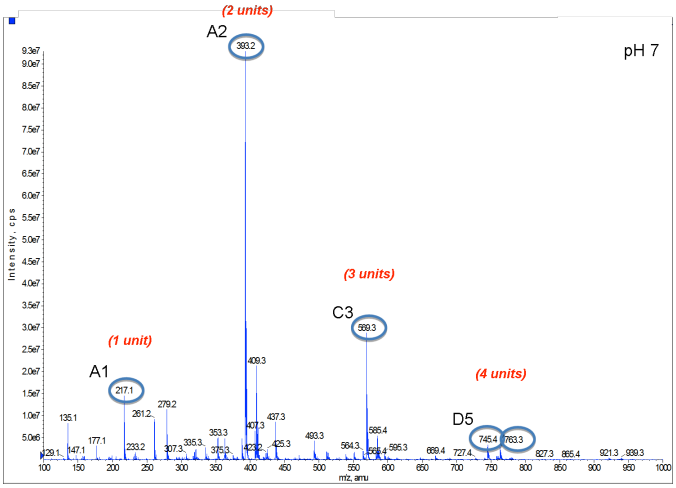


Fig. 50. MS data of GPTMS reaction at pH 7 after 24 h.

Finally, under strong basic conditions (pH 11) and as also assessed by the NMR experiments, GPTMS showed a slow hydrolysis rate of methoxy groups. In fact, the MS spectra recorded after 30 minutes of reaction identified unreacted GPTMS **D1** (259.2 M+Na) as the main signal. Additionally, at this time it was possible to observe the formation of small amounts of the tetramethoxy dimer **A3** (Fig. 52, 449.3 M+Na) and the corresponding partially hydrolysed trimethoxy **D2** (435.4 M+Na), dimethoxy **C2** (421.3 M+Na), and monomethoxy **B2** (407.4 M+Na) dimers, together with the fully de-methoxylated **A2** dimer (393.4 M+Na). Also, the fully (**B5**, 639.4 M+Na) and partially methoxylated trimers were present, with four (**A5**, 625.4 M+Na), three (**D4**, 611.4 M+Na), two (**B4**, 597.4 M+Na), and one (**D3**, 583.4 M+Na) methoxy group and the fully hydrolysed (**C3**, 569.3 M+Na) species. In three hours, the condensation reaction evolved towards the formation of the methoxylated tetramer **B6** (829.5 M+Na). Unfortunately, after this reaction time, almost all GPTMS products formed a sticky precipitate, and additional MS spectra could not be further recorded. Thus, by integrating the NMR and MS analysis, it is evident that the behaviour of GPTMS under basic conditions drastically differs from that in acidic conditions. The hydrolysis of the methoxy groups is slower at basic pH, as witnessed by the formation of methoxylated oligomers, deeply influencing the kinetics of silicon network formation. Within the first three hours of reaction, at pH 2, the pentamer was already present, while at pH 11, only a

small amount of tetramer was detectable. Thus, the polymerization was significantly faster at acidic pH than in basic conditions. Despite that, at basic pH, the formation of a white solid was fast and predominant, while at acidic pH no precipitation was detectable even after months. The precipitation at basic pH might be ascribed to the different solubility of the products that are being formed. At acidic pH, the polymers possessed free hydroxyl groups, both on the silicate and on the organic substituent (the diol is formed from epoxide hydrolysis); while at very basic pH, the oligomers were mainly methoxylated at the trimethoxysilane end of the molecule, and the epoxide did not form the corresponding diol. Thus, the presence of different amounts of hydroxyl groups at different pH might account for the observed different solubilities. The MS data of the GPTMS reaction at slightly basic pH (pH 9) are similar to those obtained at pH 11; the main difference is in the kinetics of methoxy group hydrolysis, that was slightly faster than at higher pH. After three hours, from the MS spectra most of the methoxy groups were hydrolysed. The higher detected oligomer is the trimer (C3, Fig. 52) and not the pentamer (D6, Fig. 52), as was the case at acidic pH.

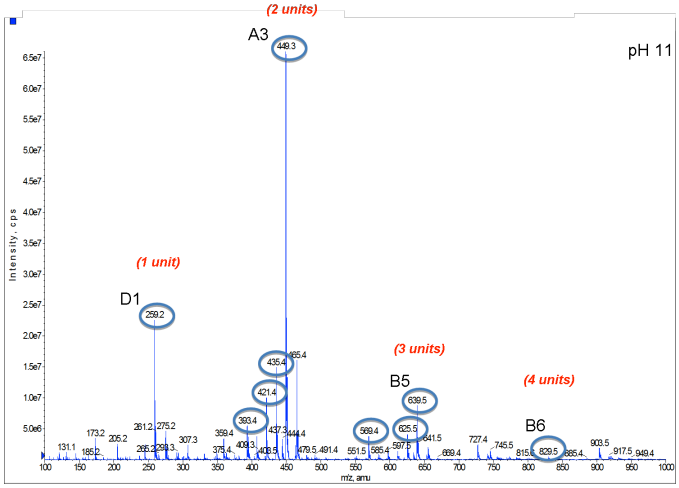
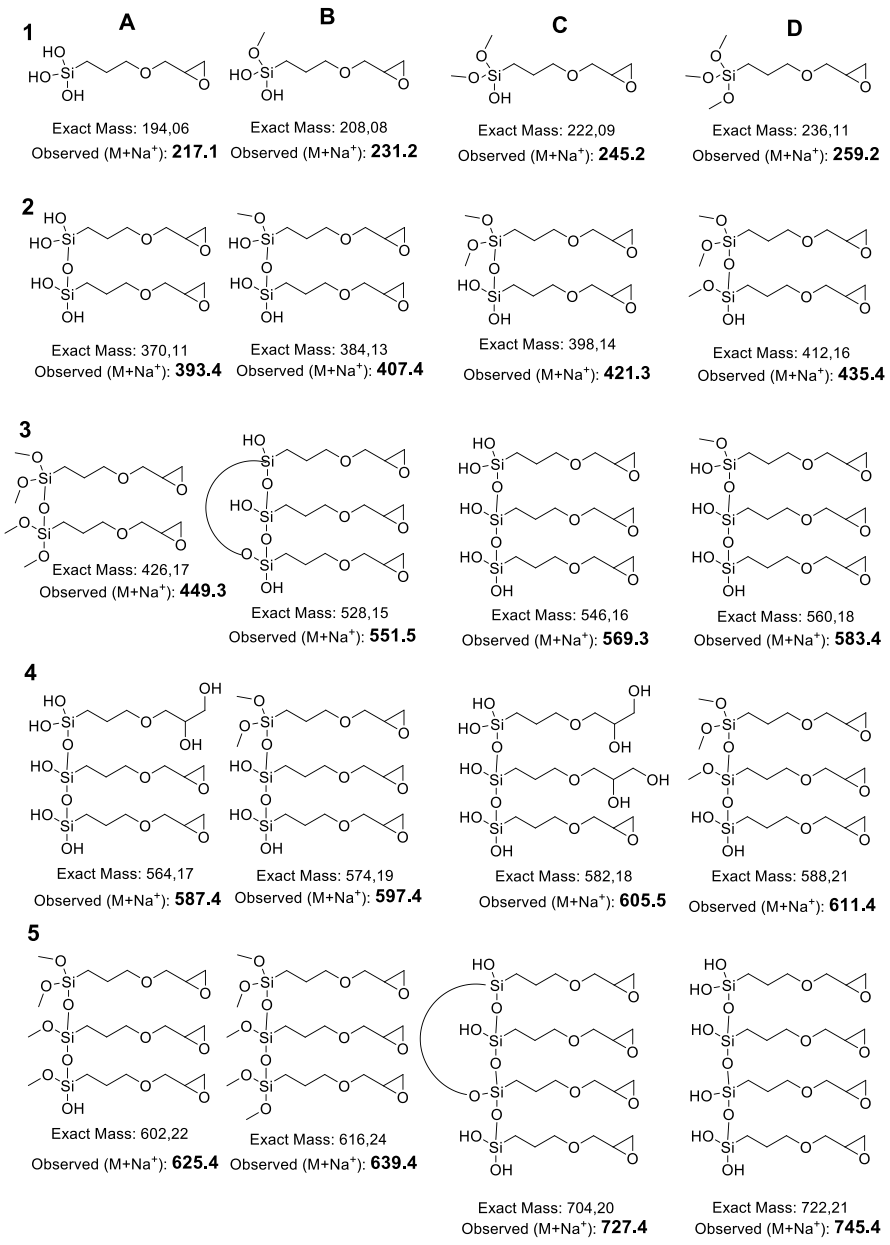


Fig. 51. MS data of GPTMS reaction at pH 11 after 3 h


A
B
C
D
1
Exact Mass: 194,06
Observed (M+Na+): 217.1
Exact Mass: 208,08
Observed (M+Na+): 231.2
Exact Mass: 222,09
Observed (M+Na+): 245.2
Exact Mass: 236,11
Observed (M+Na+): 259.2
2
Exact Mass: 370,11
Observed (M+Na+): 393.4
Exact Mass: 384,13
Observed (M+Na+): 407.4
Exact Mass: 398,14
Observed (M+Na+): 421.3
Exact Mass: 412,16
Observed (M+Na+): 435.4
3
Exact Mass: 426,17
Observed (M+Na+): 449.3
Exact Mass: 528,15
Observed (M+Na+): 551.5
Exact Mass: 546,16
Observed (M+Na+): 569.3
Exact Mass: 560,18
Observed (M+Na+): 583.4
4
Exact Mass: 564,17
Observed (M+Na+): 587.4
Exact Mass: 574,19
Observed (M+Na+): 597.4
Exact Mass: 582,18
Observed (M+Na+): 605.5
Exact Mass: 588,21
Observed (M+Na+): 611.4
5
Exact Mass: 602,22
Observed (M+Na+): 625.4
Exact Mass: 616,24
Observed (M+Na+): 639.4
Exact Mass: 704,20
Observed (M+Na+): 727.4
Exact Mass: 722,21
Observed (M+Na+): 745.4

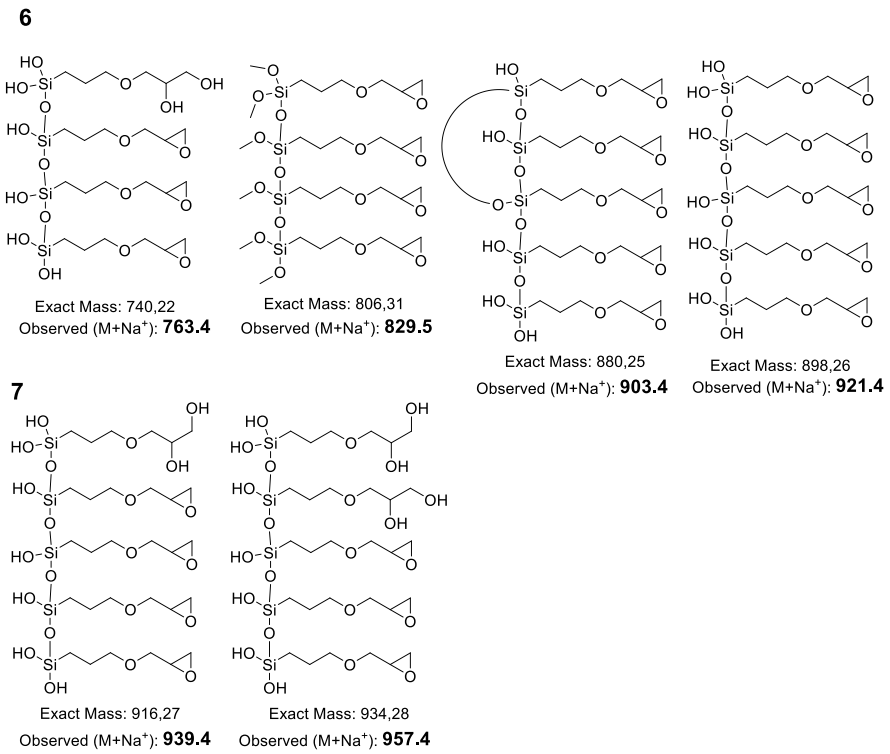


Fig. 52. Structure and mass values of the products found in the reaction system at different pH. Compounds are identified according to their location in the matrix.

By the synergy between NMR and mass spectroscopy the reactivity of GPTMS at different pH values was investigated in detail. GPTMS is one of the most used organosilanes for the fabrication of hybrid biomaterials to form covalent bonds between the organic and inorganic networks. Both spectroscopic techniques must be integrated in order to have a complete picture of the system; MS data elucidate very clearly the degree of the condensation reaction, which is not observable by NMR that retrieve slightly broad-shaped

signals of condensed species, without any indication of the dimensions of the polymers being formed.

Combining the NMR and MS results led to the conclusions that:

- At very acidic pH (pH 2) silanols form and condensation begins immediately after they form. The epoxide then hydrolyses, forming the corresponding diols on higher oligomers
- At neutral pH, all reactions are very slow: methoxy hydrolysis is slow and the condensation reaction occurs on partially de-methoxylated monomers. The oxirane ring is hydrolysed later, on the tetramers and higher oligomers.
- At very basic pH (pH 11) the hydrolysis of the methoxyl groups is slow, and condensation occurs mainly among mone-demthoxylated units. The epoxide ring does not open at all.

These results lead to some important considerations:

1. In order to functionalise polymers with GPTMS via nucleophilic attack, slightly acidic conditions (i.e. pH 5) are needed, since at neutral pH the catalysis will be too slow; while increasing too much the acidity hydrolysis of the epoxide to the corresponding diol can be the prevalent reaction, limiting the attack by the title nucleophile. Lengthy reaction times may be needed for the reaction to go to completion.

2. Under basic conditions, the precipitation process is the dominant factor, and the opening of the epoxide does not occur.
3. pH and reaction time are therefore critical in hybrid synthesis.

Finally, it should be noted that we used Mass spectroscopy for the first time for the characterisation of GPTMS reaction system.

7.1. API Conclusion

In order to elucidate API features and to block the LPS biosynthesis in gram-negative bacteria, I synthesized 16 arabinose-5-phosphate analogues. The interaction of the synthesized analogues with the enzyme API was studied by STD-NMR experiments and the effect on gram-negative bacterial growth was assayed on different strains of *E. Coli* and *P. Aureginosa*.

Previous NMR studies explored the interaction between API and A5P analogues deoxy in position 3 or 4. Thus 3-deoxy-D-arabinose-5-phosphate (**11**) and 4-deoxy-D-arabinose-5-phosphate (**22**) were synthesized and their biological activity was tested, but none of them behaved as inhibitor.

A set of cyclic A5P analogues was synthesized in order to study the ability of KdsD from *P. aeruginosa* to recognize substrates in their cyclic form. This information will be useful for the design and optimization of more potent inhibitors.

The arabinose derivatives synthesised in this work allowed us both to evaluate the ability of the enzyme to bind the substrate in their furanosidic form and to test the efficiency of the carboxylate group to mimic properly the natural phosphate group. At the same time we tested *in vivo* the activity of these derivatives on a panel of Gram-negative bacterial strains.

None of the arabinose derivatives exerted inhibitory activity against wild type strains not even when the hexose phosphate uptake (*uhp*)

system was induced, while a low inhibitory effect was observed against the *E. Coli* AS19 mutant cells, which have a defective outer membrane. These results suggest that an intact outer membrane is the major barrier for the entry of compounds **23-26** into the cell, as these compounds can permeate, in AS19 cells even when the *uhp* system is not induced. Based on NMR epitope mapping data, KdsD is the target protein of the inhibitory activity of compounds **23-24** suggesting that KdsD is able to accommodate the substrate in its cyclic form. Moreover, the cyclic compounds **23** and **24** are active against permeable *E. coli* cells, thus targeting not only KdsD but also GutQ,an additional API function in E. coli capable to substitute KdsD in Kdo synthesis. Therefore our compounds are likely to be API inhibitors. Compounds **25** and **26** did not interact with KdsD in epitope mapping studies, suggesting that the carboxylate group is not a good mimic of the phosphate group. However, the biological activity of these compounds against AS19 mutant suggests that another essential protein must be targeted. A candidate target protein of these compounds is the Kdo-8P synthase, indeed the chemical structure of compounds **25** and **26** suggests a possible function as phosphoenolpyruvate mimetics. Epitope mapping studies using the purified Kdo-8P syntase are needed to verify this hypothesis.

The synthesized compounds represent two new classes of potential inhibitors of the Kdo biosynthetic pathway based on arabinose-5-phosphate cyclic structure acting on different target enzymes.

Nevertheless, some questions remain to solve and additional rounds of optimization are still needed on these compounds to improve membrane permeability and target specificity in order to obtain promising drug leads.

The phosphate group is fundamental for the enzyme recognition, but due to its bio-instability it is important to find a good mimetic. Thus a small library of A5P analogues with different phosphate mimetics was synthesised. The ability of KdsD from *P. aeruginosa* to recognize these phosphate mimetics was explored by STD-NMR. Based on NMR epitope mapping data, KdsD is able to recognise only phosphate isosteric groups, such as the methylphosphonate **31** and the difluoromethylphosphonate **36**. However both compounds bind the enzyme weaker than the natural substrates. Further studies need to be done, in order to confirm the ability of API to isomerise compounds **31**, **36**. In addition *in vivo* assays need to be performed in order to assess if compounds **31**, **34-41** can interfere on the Kdo biosynthesis, thus acting as antibiotics agents.

7.2. Akt Conclusion

In order to inhibit the PH domain in Akt, compound **76** was synthesized. This phosphatidil inositol mimetic is part of a novel class of kinase inhibitors, based on a glucose scaffold straightforwardly obtained from the commercially available 2,3,4,6-tetra-O-acetyl-α-D-glucopyranosyl bromide in only four synthetic steps. Preliminary biological data on dendritic cell and cardiomyocytes systems identified compound **76** as a promising lead. In addition, the abolition of compound **76** effects by Akt1 silencing provides further evidence of the ability of the compound to exert biological effects through the modulation of PH-domain activated signalling. However when sensitivity of tumour cells to **76** was examined on the human ovarian carcinoma cell line IGROV-1 indicated that **76** displayed marginal inhibitory effects when tested at 50 μM. Thus lead optimization is required and in particular synthesis and studies of new sulphonate and fluorinated analogues is underway.

7.3. CI-M6PR Conclusion

Two new mannose-6-phosphate analogues have been synthesised. α–propargyl D-mannose-6-phosphate (93) was synthesized in only three steps starting from D-mannose, using a new phosphorylation method selective for primary hydroxyl groups. α–propargyl D-mannose-6-difluoromethylphosphonate (82) was synthesised in a 10-steps procedure (2g scale) and it is a new mimetic of M6P. Anyway in the last step there are still purification's problems to be solved. As already explained, the use of an ion exchange resin did not give good results, so would be good to try a fast silica filtration followed by HPLC purification.

L-Homoazidoalanine was synthesised as well and used for the production of four Np276 proteins modified with Haa starting from amino acid 41.

Np276 plasmids mutated in the desired positions in order to get single –15 – 20 – 25 – 30 Å rulers (and other "intermediate" plasmids) were successfully produced.

I found the conditions for the overexpression and purification of the single – 20 – 25 – 30 Å rulers.

Finally I found the right conditions to successfully glycosylate azido-Np276 with the synthesised phosphate (93). This was the first time that a protein was glycosylated by Huigens cycloaddition with a deprotected phosphorylated carbohydrate.

In future we should functionalize the N-terminal end of the protein with a suitable fluorescent probe. Once we obtain all the protein rulers modified with the fluorescent probe and the mannose 6-phosphate, or the mannose 6-difluorophosphonate, we could test them on hepatic stellate cells, the main target for cystic fibrosis.

The fluorescent probe will make easier and more elegant the biological evaluation of the compounds, because by microscope it will be possible to check the cell's uptake of the modified proteins.

I expect the highest uptake for the ruler with the right distance to have a bivalent binding with the receptor, and the lowest uptake for the single ruler, which for definition can't have a bivalent binding.

With these experiments we can reach different important goals:

1. to show the use of a protein as molecular ruler.
2. to prove the bivalent binding of CIM6PR, finding the right inter-phosphate distance. If this will happen we will have the key to show as many important biological aspects, because we already know that the bivalent binding increase the lysosomal uptake of important factors for the cells.
3. to prove that for CIM6PR, the difluorophosphonate can be a good mimetic of the phosphate. For many medical applications new M6P mimetics are required, and this one could be very promising, not only because is not hydrolysable, but also because of the presence of fluorine atoms could open future MRI applications.

7.4. NMR & biomaterials: Conclusion

The synthesis of "hybrid" materials is a multistep process, involving inorganic and organic polymerization; only little information is available on the reactivity of these hybrid systems. Thus, knowledge of the system's reactivity is required in order to tune the chemical, physical and biological property of the material. For the first time the formation of a hybrid material (GPTMS/diamino PEG) has been studied in detail by solution-phase NMR and this spectroscopy tecnique has been extremely useful to characterise the reaction features of the bio-glass fabrication.

In addition the synergy between NMR and mass spectroscopy was very usefull for studying the reactivity of 3-glycidoxypropyltimethoxysilane (GPTMS), one of the most common silane used for hybrid biomaterial synthesis. In particular the behaviour of the organic and inorganic components, as well as their interplay, was elucidate as a function of pH.

8. Experimental Part

8.1. API

General methods

Solvents were dried over molecular sieves, for at least 24 h prior to use, when required. When dry conditions were required, the reaction was performed under Ar or N_2 atmosphere. Thin-layer chromatography (TLC) was performed on silica gel $60F_{254}$ coated glass plates (Merck) with UV detection when possible, charring with a conc. H_2SO_4/EtOH/H_2O solution (10:45:45 v/v/v), or with a solution of $(NH_4)_6Mo_7O_{24}$ (21 g), $Ce(SO_4)_2$ (1 g), conc. H_2SO_4 (31 mL) in water (500 mL) and then heating to 110°C for 5 min. Flash column chromatography was performed on silica gel 230-400 mesh (Merck). Platinum oxide was filtered by Acrodis® Premium 25 mm Syringe Filter, with 0.45 um Nylon membrane. Routine ^1H and ^{13}C NMR spectra were recorded at 400 MHz (^1H) and at 100.57 MHz (^{13}C). Chemical shifts are reported in parts per million downfield from TMS as an internal standard; *J* values are given in Hz. Mass spectra were recorded on a System Applied Biosystems MDS SCIEX instrument (Q TRAP, LC/MS/MS, turbon ion spray) or on a System Applied Biosystem MDS SCIEX instrument (Q STAR elite nanospray). ESI full MS were recorded on a Thermo LTQ instrument by direct inlet; relative percentages are shown in brackets.

STD-NMR experiments *were performed by dott. C. Airoldi, Biotecnology department, University of Milano-Bicocca.*

All of the experiments were recorded with a Varian Mercury (400 MHz) or a Bruker Advance III (600 MHz) instrument with a cryoprobe. The ligand resonances were assigned by $^1H,^1H$ COSY and $^1H,^{13}C$-HSQC NMR spectroscopy. A basic 1D-STD sequence was used with the on-resonance frequency of 7.9 ppm and the off-resonance frequency of 40 ppm. A train of Gaussian-shaped pulses of 50 ms each was employed, with total saturation times of the protein envelope ranging from 5.0 to 0.3 s. The total saturation time was adjusted by the number of shaped pulses. A $T_{1\rho}$ filter of 2 ms was employed to eliminate the background signals from the protein. All of the samples were dissolved in deuterated PBS (pH 7.0) at 25 or 37°

C. Total sample volumes were 550 mL. The on- and off-resonance spectra were acquired simultaneously with the same number of scans. The STD spectrum was obtained by subtraction of the on-resonance spectrum from the off-resonance spectrum. Subtraction was performed by phase cycling to minimise artefacts arising from magnet and temperature instabilities. Reference experiments with samples containing only the free compounds tested were performed under the same experimental conditions to verify true ligand binding. The effects observed in the presence of the protein were due to true saturation transfer, because no signals were present in the STD spectra obtained in the reference experiments, except for residues from HDO, indicating that artefacts from the subtraction of compound signals were negligible.

Biological assays *were performed by Prof. A. Polissi, Biotecnology department, University of Milano-Bicocca.*

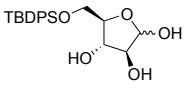
Expression and purification of Pa-KdsD:

The *P. aeruginosa*KdsD (Pa-KdsD, PA4457) was expressed by pQE31/Pa-KdsDplasmid that encodes Pa-KdsD with an N-terminal hexahistidine tag[246]. Cultures of *E. coli* M15/pREP4 harboring pQE31/Pa-KdsD plasmid were grown overnight in LD broth (Sabbatini*et al.*, 1995) supplemented with ampicillin (100 μg mL^{-1}) and kanamycin (25 μg mL^{-1}), diluted 1:100 in fresh medium (1 L)and grown up to mid-logarithmic phase (OD_{600}, 0.6) at 37 °C with shaking. Induction of Pa-KdsDwas then carried out for 4 h with IPTG (0.5 mM). Cells were harvested, washed with sodium phosphate (20 mM, pH 7.0) and re-suspended in sodium phosphate (50 mM, pH 8.0), NaCl (300 mM), imidazole (10 mM) supplemented with DNase (0, 1 mg mL^{-1}) and phenylmethylsulfonyl fluoride (1 mM). After treatment with lysozyme (1 mg mL^{-1}) for 30 min, cells were disrupted by a single cycle through a Cell Disruptor (One Shot Model by Constant Systems LTD) at 25,000 psiand unbroken cells were removed by centrifugation at 39,000g for 30 min. The supernatant was loaded at a flow rate of 0.5 mL min^{-1} onto a Ni-NTA Agarose (Qiagen) column (bed volume 5 mL) pre-equilibrated with sodium phosphate (50 mM, pH 8.0) and NaCl (300 mM). The column was washed with 10 volumes ofsodium phosphate (50 mM, pH 8.0), NaCl (300 mM) and imidazole (20mM) and the enzyme was eluted with a 5 mL step-wise imidazole gradient (100 to 300 mM)

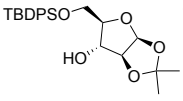
in the same buffer. Fractions (5 mL) were collected and those displaying the highest purity (>95%, as assessed by SDS-PAGE analysis) were dialyzed in sodium phosphate (50 mM, pH 8.0). The isomerase activity of the protein was assessed as described previously (Airoldi*et al.*, 2011). Protein aliquots (600 µg each) were lyophilized and stored at - 80 °C.

MIC determination

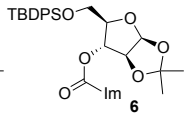
Minimal inhibitory concentrations (MIC) were determined following standard protocols for testing susceptibility to antibiotic agents[247]. The bacterial strains used were *Pseudomonas aeruginosa* PAO1 (ATCC 15692), *Escherichia coli* K12 wild type strain BW25113 [248], and the permeable *E. coli*mutant strain AS19 [123]. The procedure for a broth microdilution test was followed. Briefly, 96-well microtiter plates were prepared with compound concentrations ranging from 2 mMto 2µM, in triplicates. For *E. coli* BW25113 the procedure was performed in presence or absence of 10 µM glucose 6-phosphate. The assayed strains were grown on Mueller-Hinton broth (MHB) and inoculums of approximately 5 x 10^5cfumL^{-1} were added to each well. The microtiter plates were incubated for 18-20 h at 37°C and the MIC of each compound determined. MIC was defined as the lowest concentration inhibitingbacterial growth.



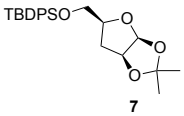
4



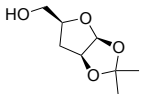
5



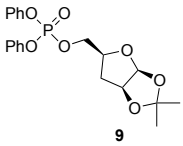
6



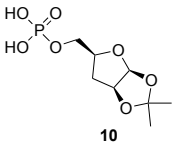
7



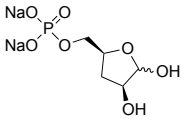
8



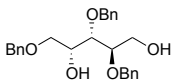
9



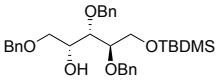
10



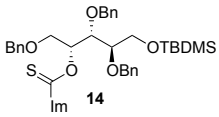
11



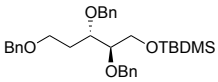
12



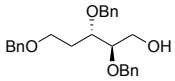
13



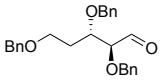
14



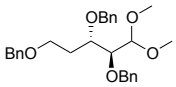
15



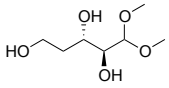
16



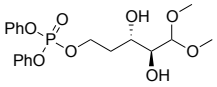
17



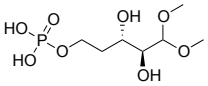
18



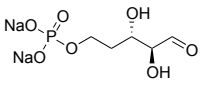
19



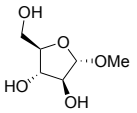
20



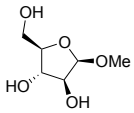
21



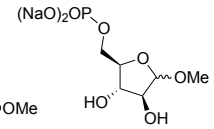
22



27

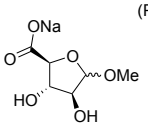


28



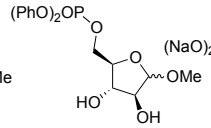
23 α-**methyl**

24 β-**methyl**



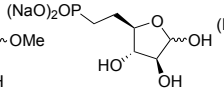
25 α-**methyl**

26 β-**methyl**

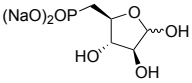


29 α-**methyl**

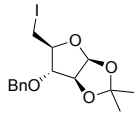
30 β-**methyl**



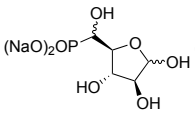
31



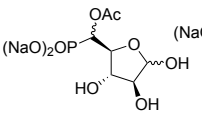
32



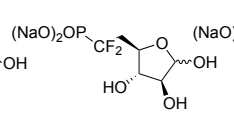
33



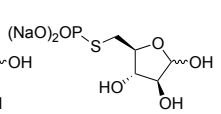
34



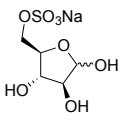
35



36



37



38

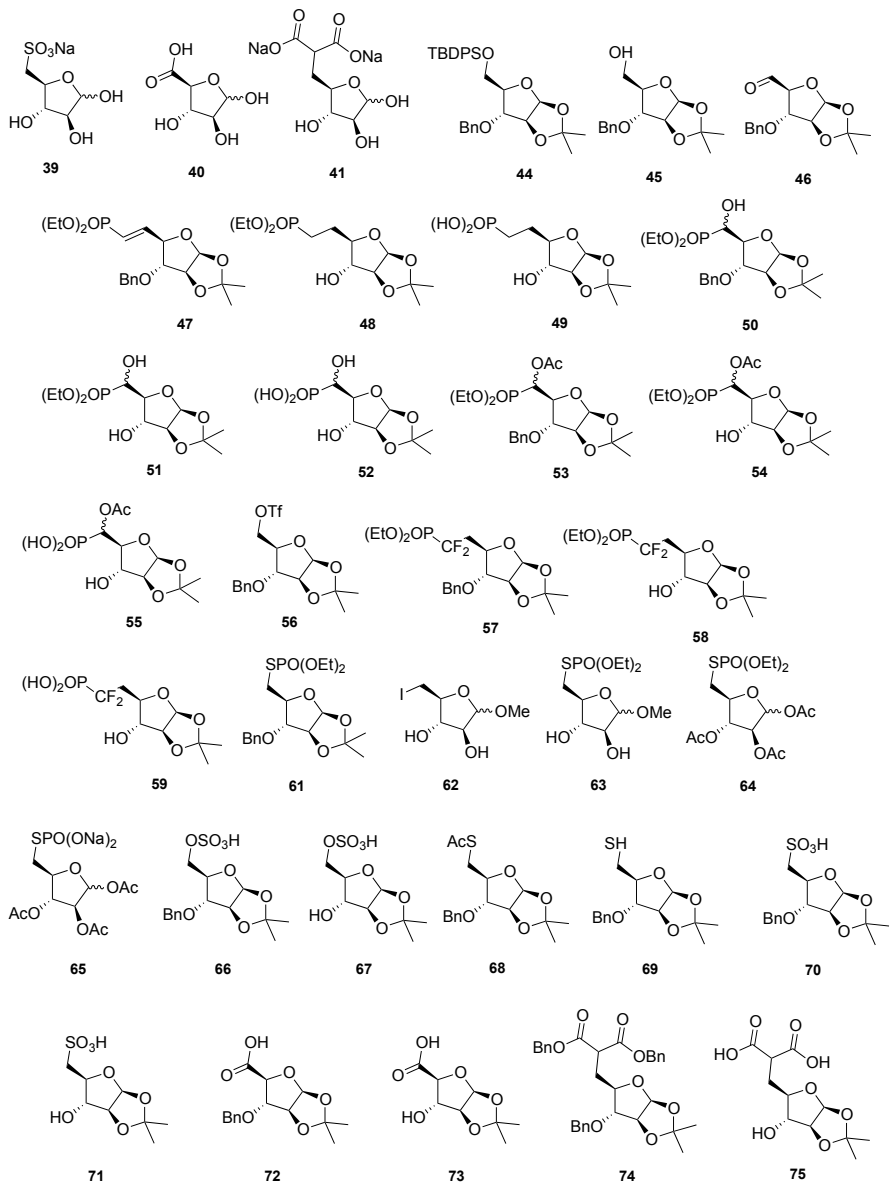


Fig. 53. Summary of the synthesised compounds.

5-O-t-Butyldiphenylsilyl-D-arabinofuranose (4)

To a stirred suspension of D-arabinose (1.83 g, 12.24 mmol, 1 eq) in dry Py (30 mL), TBDPSC1 (3.83 mL, 14.9 mmol, 1.2 eq) was added dropwise at 0◦C under argon atmosphere. Afer 10 min the reaction was stirred at 4◦C overnight. EtOH was added to quench the reaction and the mixture was stirred for 15 min. The mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (5:5, PE:EtOAc) affording **4** (3.547 g, 75% yield) as a yellow oil (mixture of α- and β- anomers). NMR and mass data are in agreement with those reported in literature.[122]

5-O-t-Butyldiphenylsilyl-1,2-O-isopropylidene-β-D-arabinofuranose (5)

To a stirred solution of **4** (3.216 g, 8.278 mmol, 1 eq) in dry CH_2Cl_2 (190 mL), 2,2-dimethoxypropane (2.18 mL, 17.79 mmol, 2 eq) and camphorsulphonic acid (0.3 g, 0.828 mmol, 0.1 eq) were added at 4◦C under argon atmosphere. The reaction mixture was stirred overnight at 4◦C. After 14 h, Et_3N was added to neutralize the reaction and the mixture was stirred for 15 min. The mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (8:2, petroleum ether:EtOAc) giving **5** (2.69 g, 76% yield) as a yellow oil. NMR and mass data are in agreement with those reported in literature.[122]

5-O-t-Butyldiphenylsilyl-3-O-(imidazolylthiocarbonyl)-1,2–O-isopropylidene-β-D-threopentofuranose (6)

To a stirred solution of **5** (1.69 g, 3.926 mmol, 1 eq) in dry dimetoxyethane (56 mL), 1,1-thiocarbonyldiimidazole (1.07 g, 7.71 mmol, 2 eq) was added under argon atmosphere. The reaction mixture was heated at reflux. After 12 hours, the reaction mixture was cooled to r t and the mixture was concentrated under reduced pressure. The residue was dissolved in EtOAc and the organic layer was washed with water. NMR and mass data are in agreement with those reported in literature.[122]

5-O-t-Butyldiphenylsilyl-3-deoxy-1,2-O-isopyropylidene-β-D-threopentofuranose (7)

To a stirred solution of crude **6** in dry toluene (78 mL), tris(trimethylsilyl) silane (782 μL, 4.645 mmol, 1.2 eq) and α,α-azobisbutyronitrile (190 mg, 1.161 mmol, 0.3 eq) were added and the resulting solution was heated and stirred at 110°C. Synthesis of 3-deoxy-*D*-threopentofuranose 5-phosphate under argon atmosphere and refluxed. After 3 hours, the mixture was cooled to rt and it was concentrated under reduced pressure. The product was purified by flash column chromatography (9:1, PE:EtOAc) affording **7** (1.15 g, 72% yield over 2 steps) as a yellow oil. NMR and mass data are in agreement with those reported in literature.[122]

3-Deoxy 1,2-O-isopropylidene-β-D-threopentofuranose (8)

To a stirred solution of **7** (1.95 g, 4.758 mmol, 1 eq) in dry THF (4.8 mL), tetrabutylammonium fluoride 1M in THF (4.8 mL, 4.8 mmol, 1 eq) was added at rt under argon atmosphere, After 1 h, silica was added and the mixture was conecentrated under reduced pressure. The product was purified by flash column chromatography (3:7 EP:EtOAc) giving **8** (712 mg 87% yield) as a yellow oil. NMR and mass data are in agreement with those reported in literature.[122]

3-Deoxy-1,2-O-isopropylidene-β-D-threopentofuranose 5-diphenylphosphate (9)

To a stirred solution of **8** (162 mg, 0.931 mmol) in dry Py (10 mL), diphenylphosphochloridate (310 μL, 1.4 mmol, 1.5 eq) was added at 0°C under argon atmosphere. After 1.5 hours EtOH was added to quench the reaction and the mixture was stirred for 15 min. The mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (6:4, petroleum ether:EtOAc) affording **9** (281 mg, 75% yield) as a white solid. NMR and mass data are in agreement with those reported in literature.[122]

3-Deoxy-1,2-O-isopropylidene-β-D-threopentofuranose 5-phosphate (10)

Compound **7** (30 mg, 0.074 mmol) was dissolved in degassed MeOH (3 ml) and PtO_2 was added (16.8 mg). The reaction mixture was stirred under H_2 at rt overnight. The catalyst was removed by syringe filter and the filtrate was concentrated under reduced pressure, obtaining **10** (19 mg, quantitative yield) as a

white solid. NMR and mass data are in agreement with those reported in literature.[122]

3-Deoxy-D-threopentofuranose 5-phosphate (11)

Compound **10** was dissolved in H_2O and the solution was transferred into an NMR-tube. The reaction was monitored by 1H NMR. After 1 h pH was neutralized by adding NaOH 1.0 M and the solution was lyophilized to give **11** in quantitative yield (mixture of α- and β- anomers). NMR and mass data are in agreement with those reported in literature.[122]

Diol (12)

To a 0°C solution of **2,3,5-tri-O-benzyl-D-arabinose** (1.445 g, 3.44 mmol, 1 eq) in ethanol (10 ml), sodium borohydride was added (170 mg, 4.4 mmol, 1.3 eq). The solution was stirred for 3 hours at room temperature, for 1 hour. The solvent was evaporated and the then the mixture was partitioned between EtOAc:HCl 5%. The organic layer was washed sequentially with water and brine, dried over Na_2SO_4 and concentrated under vacuum affording the open-chain diol (1.43 g, quantitative yields) as a viscous oil NMR and mass data are in agreement with those reported in literature.[121]

Sylil ether (13)

A solution of the diol **12** (2.187 g, 5.18 mmol, 1 eq) in dry pyridine (12 ml) was stirred at 0°C and t-butyldimethylsilyl chloride (1.01 g, 6.734 mmol, 1.3 eq) was added. The reaction was allowed to warm to room temperature over 2 h and was quenched adding EtOH, then it was concentrated under reduced pressure. The residue was purified by flash chromatography, PE:EtOAc 7:3 to give the silyl ether **13** (2.71 g, 96. NMR and mass data are in agreement with those reported in literature.[121]

Thiocarbonylimidazole ester (14)

To a solution of silyl ether **13** (1.73 g, 3.22 mmol, 1 eq) in dry tetrahydrofuran (36 ml) was added 1,l'-thiocarbonyldiimidazole (1.47 g, 8.26 mmol, 2.6 eq) and the

solution was heated at reflux for 5 hours. After cooling to room temperature, the solution was concentrated under reduced pressure. Purification of the residue by chromatography on silica EP:EtOAc 8.8:1.2 gave the thiocarbonylimidazole ester (1.474 g, 75%). NMR and mass data are in agreement with those reported in literature.[121]

Deoxysilyl ether (15)

A solution of thiocarbonylimidazole ester **14** (105 mg, 0.163 mmol, 1 eq) from the above, in dry toluene (0.8 ml), was added to a refluxed tributyltin hydride (70 ml, 0.26 mmol, 1.6 eq) in toluene (2 ml). After 6 h, the reaction mixture was cooled to room temperature, the solvent was concentrated to dryness and the product was purified by flash chromatography EP:EtOAc (9.5:0.5) to give deoxysilyl ether **15** (2.25 g, 74%) as a colorless oil. NMR and mass data are in agreement with those reported in literature.[121]

Alcohol (16)

To a stirred solution of **15** (340 mg, 0.6 mmol, 1eq) in dry THF (4 mL), tetrabutyl ammonium fluoride (1 M THF solution, 2 ml, 2 mmol, 3 eq) was added. The reaction mixture was stirred for 1 hour and then the mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (7:3 PE:EtOAc) giving **16** (258 mg, 70% yield) as a white solid. NMR and mass data are in agreement with those reported in literature.[121]

Aldehyde (17)

To a stirred solution of of alcohol from the above (331 mg, 0.81 mmol, 1eq) in dry DCM (10 mL), Dess-Martin reagent (414 mg, 0.97 mmol, 1.2 eq) was added. The reaction mixture was stirred for 30 minutes and then it was neutralized adding $NaHCO_3$ saturated solution and one volume of $Na_2S_2O_3$ saturated solution. The mixture was extracted 3 times with DMC; the organic phase is dried with Na_2SO_4, filtered and concentrated under reduced pressure and immediately used for the next reaction without any characterization. NMR and mass data are in agreement with those reported in literature.[121]

Dimethylacetal (18)

To the fresh solution of **17** from the above with methanol (0.22 ml) and trimethyl orthoformate (2.2 ml) at O°C, was added concentrated sulfuric acid (4 mL). The solution was allowed to warm to room temperature over 45 min, then quenched with 1.0 M $NaHCO_3$. The reaction mixture was partitioned between ethyl acetate and brine, and the organic layer was dried over anhydrous $NaSO_4$, and concentrated under reduced pressure. Purification of the residue by flash chromatography 9:1 EP:EtOAc gave the dimethylacetal (328 mg, 90% over two steps). NMR and mass data are in agreement with those reported in literature.[121]

Triol (19)

The mixture of the dimethylacetal **18** (84 mg, 0.186 mmol, 1 eq) and $Pd(OH)_2$ (80 mg) in EtOAc/EtOH 1/1 (2.4 ml), was stirred under vacuum in order to remove the residual gas and then H_2 was added. The reaction was stirred at room temperature for 24 hours and then the catalyser was filtered on a syringe filter and washed with EtOH. The solvent was evaporated giving **19** as colourless oil (34 mg, 99% yield). NMR and mass data are in agreement with those reported in literature.[121]

Diphenylphosphate (20)

To a solution of **19** (378 mg, 2.2 mmol, 1eq) in dry pyridine (5ml) was added diphenylphosphochloridate (709 mg, 2.64 mmol, 1.2 eq) and the solution was stirred 3 h at 0°C. The reaction was allowed to warm to room temperature over 1 hour then it was quenched with ethanol. After removal of solvents, the residue was purified by chromatography on EP:EtOAc 2:8 to give the protected phosphate **20** (542 mg, 60%). NMR and mass data are in agreement with those reported in literature.[121]

Phosphate (21)

A mixture of protected **20** (49 mg, 0.119 mmol) and PtO_2, (40 mg) in methanol (2 ml) was stirred under H_2 overnight. The catalyst was removed by filtration through syringe filter. The solvent was evaporated affording **21** in quantitative yield. NMR and mass data are in agreement with those reported in literature.[121]

4-deoxy-5-phosphate-D-arabinose (22)

The phosphate **21** from above was dissolved in D_2O (0.7 mL) and the hydrolysis of the methoxy groups was followed by ^{1}H-NMR. Then the pH was neutralized with NaOH 5% and the sample was freeze-dried giving **22** as sodium salt in quantitative yields. NMR and mass data are in agreement with those reported in literature.[121]

Methyl D-arabinofuranoside (27, 28)

To a solution of D-arabinose (1 eq) in MeOH (30 mL/g D-ara), AcCl (100 uL/g D-ara) was added and the reaction was stirred at room temperature for 18 h. The reaction was quenched by the addition of TEA and concentrated by rotavapor. The resulting oil was purified by flash chromatography (EtOAc-MeOH, 9.5:0.5) and the two anomers were separated to give the pure methyl furanosides (970 mg, 89%, a/b 1:0.39).

α-methyl D-arabinofuranoside (27)

^{1}H NMR (400 MHz, CD_3OD) δ 4.68 (d, $J_{1,2}$= 1.4 Hz, 1H, H1), 3.85 (dd, $J_{2,3}$ = 3.5, $J_{2,1}$= 1.4 Hz, 1H, H2), 3.84-3.80 (m, 1H, H4), 3.75 (dd, $J_{3,4}$= 6.2, $J_{3,2}$= 3.5 Hz, 1H, H3), 3.66 (dd, $J_{5a,5b}$= 11.9, $J_{5a,4}$=3.2 Hz, 1H, H5a), 3.55 (dd, $J_{5b,5a}$= 11.9, $J_{5b,4}$= 5.3 Hz, 1H, H5b), 3.28 (s, 3H, OCH_3).

^{13}C NMR (101 MHz, CD_3OD) d 111.25 (C-1), 86.18 (C-4), 84.07 (C-2), 79.40 (C-3), 63.77 (C-5), 53.9 (OCH$_3$).

MS: m/z = 187.1 [M+Na]

β-methyl D-arabinofuranoside (28)

^{1}H NMR (400 MHz, CD_3OD) δ 4.63 (d, $J_{1,2}$ = 4.4 Hz, 1H, H1), 3.85 (dd, $J_{2,3}$= 7.6, $J_{2,1}$= 4.4 Hz, 1H, H2), 3.79 (t, $J_{3,4}$= 7.1 Hz, 1H, H3), 3.94-3.63 (m, 1H, H4), 3.57 (dd, $J_{5a,5b}$= 11.6, $J_{5a,4}$= 3.6 Hz, 1H, H5a), 3.44 (dd, $J_{5b,5a}$ = 11.6, $J_{5b,4}$=7.2 Hz, 1H, H5b), 3.30 (s, 3H, OCH_3).

^{13}C NMR (101 MHz, CD_3OD) δ 104.44 (C-1), 84.87 (C-4), 79.45 (C-3), 77.25 (C-2), 65.99 (C-5), 55.96 (OCH$_3$).

MS: m/z = 187.1 [M+Na]

Methyl D-arabinuronic acid ammonium salt (25/26)

Methyl D-arabinofuranoside (**27**/**28**) (1 eq) was dissolved in a saturated aqueous solution of sodium bicarbonate (0.2 M of **27**/**28**). A 0.5 M aqueous solution of potassium bromide (0.3 eq) and a 0.1 M acetonitrile solution of TEMPO (.5 eq) were added under stirring. A 0.35 M NaOCl solution (1.5 eq) is added slowly and the reaction is stirred for 2 hours at room temperature. The solvent is evaporated by rotavapor and the resulting solid was purified by flash chromatography (IPA-NH_4OH 33%, 8:2, v/v) to give the ammonium carboxylate.

α-Methyl D-arabinuronic acid ammonium salt (25)

By subjecting α-methyl D-arabinofuranoside (**27**) (235 mg, 1.433 mmol) to the procedure for the synthesis of methyl D-arabinuronic acid ammonium salt, compound **25** was obtained as a yellowish solid (180 mg, 64%).

^{1}H NMR (400 MHz, D_2O) δ 4.96 (d, $J_{1,2}$= 1.0 Hz, 1H, H1), 4.26 (d, $J_{4,3}$= 4.3 Hz, 1H, H4), 4.09 (dd, $J_{3,4}$= 4.3, $J_{3,2}$=2.3 Hz, 1H, H3), 3.96 (dd, $J_{2,3}$= 2.3, $J_{2,1}$=1.0 Hz, 1H, H2), 3.37 (s, 3H, OCH_3).

^{13}C NMR (101 MHz, D_2O) δ 180.92 (C-5), 107.27 (C-1), 82.43 (C-4), 78.43 (C-2), 77.85 (C-3), 53.47 (OCH$_3$).

MS: m/z = 177.1 [M-1]

β-Methyl D-arabinuronic acid ammonium salt (26)

By subjecting β-methyl D-arabinofuranoside (**28**) (156 mg, 0.951 mmol) to the procedure for the synthesis of methyl D-arabinuronic acid ammonium salt, compound **26** was obtained as a yellowish solid (115 mg, 62%).

^{1}H NMR (400 MHz, D_2O) δ 4.88 (d, J = 4.5 Hz, 1H, H1), 4.26 – 4.19 (m, 1H, H3), 4.10 (d, J = 6.5 Hz, 1H, H4), 4.04 (dd, J = 7.5, 4.5 Hz, 1H, H2), 3.43 (s, 3H, OCH_3).

^{13}C NMR (101 MHz, D_2O) δ 177.18 (C-5), 102.54 (C-1), 80.15 (C-4), 76.95 (C-3), 76.18 (C-2), 55.42 (OCH$_3$).

MS: m/z = 177.1 [M-1]

Methyl D-arabinofuranoside-5-diphenyl phosphate (29/30)

Methyl D-arabinofuranoside (**27**/**28**) (126 mg, 0.768 mmol) was dissolved in dry pyridine (8 mL). At 0°C cloro diphenylphosphate (176 uL, 0.845 mmol, 1.1 eq) was added and then the reaction is stirred at room temperature for 2 hours. The reaction was quenched adding EtOH and then it was stirred for 15 minutes before

evaporating the solvent. The resulting oil was purified by flash chromatography (EP-EtOAc, from 2:8 to 1:9, v/v) to give pure diphenyl phosphate.

α-Methyl D-arabinofuranoside-5-diphenyl phosphate (29)

By subjecting α-methyl D-arabinofuranoside (**27**) (126 mg, 0.768 mmol) to the procedure for the synthesis of methyl D-arabinofuranoside-5-diphenyl phosphate, compound **29** was obtained as a colourless oil (250 mg, 82%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.56 – 6.87 (m, 10H, HPh), 4.82 (s, 1H, H1), 4.46 – 4.31 (m, 1H, H5), 4.15 (q, $J_{4,3}$= 4.1 Hz, 1H, H4), 4.03 (d, $J_{2,1}$= 1.2 Hz, 1H, H2), 3.93 (dd, $J_{3,4}$= 4.1, 2.5 Hz, 1H, H3), 3.45 (bs, 2H, O*H*), 3.33 (s, 3H, OCH_3).

^{13}C NMR (101 MHz, $CDCl_3$) δ 150.44 (d, $J_{Cq,P}$= 3.4 Hz, C_{quat}Ph), 150.37 (d, $J_{Cq,P}$= 3.4 Hz, C_{quat}Ph), 130.01 (C-Ph), 125.73 (C-Ph), 120.21 (C-Ph), 108.97 (C-1), 83.21 (d, $J_{C4,P}$= 6.8 Hz, C-4), 80.85 (C-2), 77.52 (C-3), 68.59 (d, $J_{C5,P}$= 6.6 Hz, C-5), 55.27 (O*C*H_3).

^{31}P NMR (162 MHz; $CDCl_3$) δ -11.67 (1 P).

MS: m/z = 419.1 [M+Na], 815.3 [2M+Na]

β-Methyl D-arabinofuranoside-5-diphenyl phospate (30)

By subjecting β-methyl D-arabinofuranoside (**28**) (126 mg, 0.768 mmol) to the procedure for the synthesis of methyl D-arabinofuranoside-5-diphenyl phosphate, compound **30** was obtained as a colourless oil (228 mg, 75%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 – 7.11 (m, 10H, HPh), 4.75 (d, $J_{1,2}$= 3.7 Hz, 1H, H1), 4.37 – 4.21 (m, 2H, H5a, H5b), 4.07 – 3.97 (m, 3H, H2, H4, H3), 3.71 (bs, 2H, OH), 3.33 (s, 3H, OCH_3).

^{13}C NMR (101 MHz, $CDCl_3$) δ 150.44 (d, $J_{Cq,P}$= 3.4 Hz, C_{quat}Ph), 150.37 (d, $J_{Cq,P}$= 3.4 Hz, C_{quat}Ph), 129.95 (C-Ph), 125.65 (C-Ph), 120.16 (C-Ph), 102.25 (C-Ph), 80.24 (d, $J_{C4,P}$= 7.4 Hz, 1C, C4), 77.81 (C-3), 75.82 (C-2), 69.74 (d, $J_{C5,P}$= 5.5 Hz, C-5), 55.37 (O*C*H_3).

^{31}P NMR (162 MHz; $CDCl_3$) δ -11.94 (1 P).

MS: m/z = 419.1 [M+Na], 815.3 [2M+Na]

Methyl D-arabinofuranoside-5-phosphate sodium salt (23/24)

Methyl D-arabinofuranoside-5-diphenyl phosphate (**29/30**) (100 mg, 0.404 mmol) was dissolved in EtOH (6 mL). Platinum oxide (92 mg, 0.404 mmol, 1 eq) was

added, the mixture was stirred under vacuum in order to remove the residual gas and then H_2 was added. The reaction is stirred at room temperature for 15 hours and then the palladium oxide was filtered on a syringe filter and washed with EtOH. The solvent was evaporated and the resulting colourless oil was dissolved in deuterated water and checked by NMR spectroscopy. The pH is neutralized by the addition of NaOH 5% and then the solution is freeze-dried, obtaining the phosphate sodium salt as a colourless solid.

α-Methyl D-arabinofuranoside-5-phosphate sodium salt (23)

By subjecting α-methyl D-arabinofuranoside-5-diphenyl phosphate (**29**) (156 mg, 0.394 mmol) to the procedure for the synthesis of methyl D-arabinofuranoside-5-phosphate sodium salt, compound **23** was obtained as colourless oil (112 mg, 99%).

^{1}H NMR (400 MHz, D_2O) δ 4.74 (d, $J_{1,2}$ = 1.3 Hz, 1H), 3.99 – 3.91 (m, 2H, H4, H5a), 3.90 – 3.79 (m, 3H, H2, H3, H5b), 3.21 (s, 3H, OCH_3).

^{13}C NMR (101 MHz, D_2O) δ 108.12 (C-1), 82.09 (d, J_{C4P}= 8.2 Hz, C4), 80.28 (C-2), 75.82 (C-3), 64.89 (d, J_{C5P}= 5.1 Hz), 54.79 (OCH_3).

^{31}P NMR (162 MHz; $CDCl_3$) δ 6.21 (1 P).

MS: m/z = 243.1 [M-1]

β-Methyl D-arabinofuranoside-5-phosphate sodium salt (24)

By subjecting β-methyl D-arabinofuranoside-5-diphenyl phosphate (**30**) (81 mg, 0.204 mmol) to the procedure for the synthesis of methyl D-arabinofuranoside-5-phosphate sodium salt, compound **24** was obtained as colourless oil (57 mg, 97%).

^{1}H NMR (400 MHz, D_2O) δ 4.76 (d, $J_{1,2}$= 4.4 Hz, 1H, H1), 4.01 (dd, $J_{2,3}$= 8.0, $J_{2,1}$= 4.4 Hz, 1H, H2), 3.96 – 3.77 (m, 4H, H3, H4, H5a, H5b), 3.27 (s, 3H, OCH_3).

^{13}C NMR (101 MHz, D_2O) δ 101.87 (C-1), 80.09 (d, J_{C4P}= 8.5 Hz, C4), 75.92 (C-3), 74.00 (C-2), 66.47 (d, $J_{C5,P}$= 5.1 Hz, C5), 54.78 (OCH_3).

^{31}P NMR (162 MHz; $CDCl_3$) δ 6.10 (1 P).

MS: m/z = 243.1 [M-1]

5-O-t-Butyldiphenylsilyl-D-arabinofuranose (42)

To a stirred suspension of D-arabinose (1.83 g, 12.24 mmol, 1 eq) in dry Py (30 mL), TBDPSC1 (3.83 mL, 14.9 mmol, 1.2 eq) was added dropwise at 0◦C under argon atmosphere. Afer 10 min the reaction was stirred at 4◦C overnight. EtOH was

added to quench the reaction and the mixture was stirred for 15 min. The mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (5:5, petroleum ether;EtOAc) affording **42** (3.547 g, 75% yield) as a yellow oil (mixture of α- and β- anomers). NMR and mass data are in agreement with those reported in literature.[122]

5-O-t-Butyldiphenylsilyl-1,2-O-isopropylidene-β-D-arabinofuranose (43)

To a stirred solution of **42** (3.216 g, 8.278 mmol, 1 eq) in dry CH_2Cl_2 (190 mL), 2,2-dimethoxypropane (2.18 mL, 17.79 mmol, 2 eq) and camphorsulphonic acid (0.3 g, 0.828 mmol, 0.1 eq) were added at 4◦C under argon atmosphere. The reaction mixture was stirred overnight at 4◦C. After 14 h, Et_3N was added to neutralize the reaction and the mixture was stirred for 15 min. The mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (8:2, petroleum ether:EtOAc) giving **43** (2.69 g, 76% yield) as a yellow oil. NMR and mass data are in agreement with those reported in literature.[122]

5-O-t-Butyldiphenylsilyl-3-O-benzyl-1,2-O-isopropylidene-β-D-arabinofuranose (44)

To a stirred solution of **43** (1.250 g, 2.919 mmol, 1eq) in dry THF (20 mL), benzyl bromide (1.84 mL, 11.676 mmol, 4 eq) and sodium hydride 60% (584 mg, 14.595 mmol, 5 eq) were added slowly in five portions. The reaction mixture was stirred for three hours and then the mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (from 100% PE to 9:1, PE:EtOAc) giving **44** (1.40 g, 93% yield) as a yellowish oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.62 – 7.54 (m, 5H, H Ph), 7.38 – 7.17 (m, 10H, H Ph), 5.82 (d, $J_{1,2}$ = 4.1 Hz, 1H, H2), 4.59 (d, $J_{2,1}$ = 4.1 Hz, 1H), 4.54 (d, J = 1.5 Hz, 2H, CH_2 Bn), 4.21 – 4.11 (m, 2H, H3, H4), 3.80 – 3.67 (m, 2H, H5), 1.27 (s, 3H, Me), 1.22 (s, 3H, Me), 0.97 (s, 9H, tBu).

^{13}C NMR (101 MHz, $CDCl_3$) δ 137.77 (1C, C_{quat} C Ph), 135.93 (1C, C Ph), 133.50 (1C, C_{quat} C Ph), 133.38 (1C, C_{quat} C Ph), 130.09 (1C, C Ph), 130.03 (1C, C Ph), 128.83 (1C, C Ph), 128.18 (1C, C Ph), 128.07 (1C, C Ph), 128.04 (1C, C Ph), 128.02 (1C, C Ph), 112.77 (1C, C_{quat}), 106.07 (1C, C1), 85.52 (1C, C2), 85,43(1C, C4),

83.15 (1C, C3), 71.99 (1C, CH_2 Bn), 63.69 (1C, C5), 27.28 (1C, Me), 27.14 (3C, Me tBu), 26.45 (1C, Me), 19.52 (1C, C_{quat} tBu).
MS: m/z = 541.4 $[M+Na]^+$, 557.4 $[M+K]^+$.

Alcohol (45)

To a stirred solution of **44** (1.29 g, 2.49 mmol, 1eq) in dry THF (20 mL), tetrabutyl ammonium fluoride (1 M THF solution, 7.47 ml, 7.47 mmol, 3 eq) was added. The reaction mixture was stirred for 1 hour and then the mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (from 7:3 to 6:4, PE:EtOAc) giving **45** (560 mg, 80% yield) as a white solid.
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.03 (m, 5H, H Ph), 5.83 (d, $J_{1,2}$ = 3.9 Hz, 1H, H1), 4.60 (d, $J_{2,1}$ = 3.9 Hz, 1H, H2), 4.58 – 4.45 (m, 2H, CH_2 Bn), 4.15 – 4.08 (m, 1H, H4), 3.89 (d, J = 2.5 Hz, 1H, H3), 3.72 – 3.59 (m, 2H, H5), 2.71 (bs, 1H, OH), 1.44 (s, 3H, Me), 1.25 (s, 3H, Me).
^{13}C NMR (101 MHz, $CDCl_3$) δ 137.39 (1C, C_{quat} C Ph), 128.73 (1C, C Ph), 128.20 (1C, C Ph), 127.98 (1C, C Ph), 113.06 (1C, C_{quat}), 105.78 (1C, C1), 85.79 (1C, C4), 85.38 (1C, C2), 82.92 (1C, C3), 72.01 (1C, CH_2 Bn), 62.79 (1C, C5), 27.29 (1C, Me), 26.52 (1C, Me).
MS: m/z = 303.4 $[M+Na]^+$, 319.3 $[M+K]^+$, 583.3 $[2M+Na]^+$.

Aldheyde (46)

To a stirred solution of **45** (100 mg, 0.357 mmol, 1 eq) in dry DCM (4 mL), Dess-Martin reagent (181 mg, 0.428 mmol, 1.2 eq) was added. The reaction mixture was stirred for 15 minutes and then it was neutralized adding $NaHCO_3$ saturated solution and one volume of $Na_2S_2O_3$ saturated solution. The mixture was extracted 3 times with DMC; the organic phase was dried with Na_2SO_4, filtered and concentrated under reduced pressure. The product was quickly filtered with a silica column (7:3, PE:EtOAc, 1% TEA) giving **46** (98 mg) as a light yellow oil that has been immediately used for the next reaction without any characterization.

Diethyl phosphonate (47)

To a solution of NaH 60% (29 mg, 0.714 mmol, 2 eq) in dry benzene (5 ml), diethyl-methyl bisphosphonate (222 ml, 0.983 mmol, 2.5 eq) was added and then the mixture was stirred at room temperature for 30 minutes. The freshly prepared aldehyde **46** (98 mg) is dissolved in dry benzene (1 + 0.25 + 0.25 mL) and added by syringe to the phosphonate solution. The reaction is stirred at room temperature for 30 minutes, then the solvent is evaporated and the resulting product is extracted with EtOAc; the organic phase is dried with Na_2SO_4, filtered and concentrated under reduced pressure. The product was purified by flash column chromatography (3.5:6.5, PE:EtOAc) giving **47** (132 mg, 90% yield in two steps) as a light yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.23 (m, 5H, H Ph), 6.86 – 6.73 (m, 1H, H6), 6.10 – 6.95 (m, 1H, H5), 5.94 (d, $J_{1,2}$= 3.8 Hz, 1H), 4.69 – 4.50 (m, 4H, H2, H4, CH_2 Bn), 4.11 – 3.98 (m, 4H, 2CH_2 OEt), 3.95 (d, J = 2.5 Hz, 1H, H3), 1.45 (s, 3H, Me), 1.32 – 1.21 (m, 9H, Me, 2CH_3 OEt).

^{13}C NMR (101 MHz, $CDCl_3$) δ 149.13 (d, $^2J_{C\text{-}P}$ = 5.7 Hz, 1C, C5), 136.81 (1C, C_{quat} Ph), 128.63 (1C, C Ph), 128.17 (1C, C Ph), 127.83 (1C, C Ph), 117.60 (d, $^1J_{C\text{-}P}$ = 188.2 Hz, 1C, C6), 113.07 (1C, C_{quat}), 106.21 (1C, C1), 85.52, 84.76, 84.50 (C2, C3, C4), 72.01(1C, CH_2 Bn), 61.75, (2C, CH_2 OEt), 26.53 (1C, Me), 26.23 (1C, Me), 16.34, (2C, CH_3 OEt).

^{31}P NMR (162 MHz; $CDCl_3$) δ 17.78 (1 P).

MS: m/z = 413.3 $[M+H]^+$, 435.4 $[M+Na]^+$, 825.4 $[2M+H]^+$, 847.4 $[2M+Na]^+$.

Alcohol (48)

Compound **47** (233 mg, 0.212 mmol) was dissolved in EtOH (6 mL). Palladium hydroxide (200 mg) was added, the mixture was stirred under vacuum in order to remove the residual gas and then H_2 was added. The reaction was stirred at room temperature for 24 hours and then the catalyser was filtered on a syringe filter and washed with EtOH. The solvent was evaporated giving **48** as colourless oil (180 mg, 98% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.85 (d, $J_{1,2}$ = 3.9 Hz, 1H, H1), 4.54 (bs, 1H, OH), 4.50 (d, $J_{2,1}$ = 3.9 Hz, 1H, H1), 4.10 – 3.98 (m, 5H, H3, 2CH_2 OEt), 3.98 – 3.91 (m,

1H, H4), 2.08 – 1.66 (m, 4H, H5, H6), 1.47 (s, 3H, Me), 1.30 – 1.22 (m, 9H, Me, $2CH_3$ OEt).

^{13}C NMR (101 MHz, $CDCl_3$) δ 112.28 (1C, C_{quat}), 105.74 (1C, C1), 87.73 (d, $^3J_{C\text{-}P}$ = 16.2 Hz, 1 C, C4), 87.09 (1C, C2), 77.97 (1C, C3), 61.84 (2C, CH_2 OEt), 26.85 (1C, Me), 26.66 (d, $^2J_{C\text{-}P}$ = 4.5 Hz, 1C, C5), 26.11 (1C, Me), 22.06 (d, $^1J_{C\text{-}P}$ = 141.8 Hz, 1C, C6), 16.44 (d, $^3J_{C\text{-}P}$ = 6.3 Hz, 2C, CH_3 OEt).

^{31}P NMR (162 MHz; $CDCl_3$) δ 32.13 (1 P).

MS: m/z = 325.3 $[M+H]^+$, 347.2 $[M+Na]^+$, 363.3 $[M+K]^+$, 671.3 $[2M+Na]^+$.

Phosphonate (49)

To a stirred solution of compound **48** (126 mg, 0.389 mmol) in dry ACN (5 mL), dry Py (315 ml, 3.885 mmol, 10 eq) and TMSBr (513 ml, 3.885 mmol, 10 eq) were added at room temperature. After 30 minutes the solution is cooled at 0°C and the reaction is quenched adding 3.4 mL of NH_4OH 33% under stirring. The ice bath is removed and the solvent is evaporated under reduced pressure. The product was purified by flash column chromatography (6:4, IPA: NH_4OH 33%) giving **49** (98 mg, 94% yield) as a white solid.

^{1}H NMR (400 MHz, CD_3OD) δ 5.73 (d, $J_{1,2}$ = 3.8 Hz, 1H, H1), 4.38 (d, $J_{2,1}$ = 3.8 Hz, 1H), 3.87 – 3.82 (m, 1H, H3), 3.81 – 3.72 (m, 1H, H4), 1.85 – 1.66 (m, 2H, H5), 1.41 (s, 3H, Me), 1.37 – 1.26 (m, 2H, H6), 1.22 (s, 3H, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 113.57 (1C, C_{quat}), 106.76 (1C, C1), 89.60 (d, $^3J_{C\text{-}P}$= 17.1 Hz, 1C, C4), 88.49 (1C, C2), 78.97 (1C, C3), 29.22 (d, $^2J_{C\text{-}P}$ = 3.8 Hz, 1C, C5), 27.32 (1C, Me), 26.84 (1C, Me), 26.16 (d, $^1J_{C\text{-}P}$ = 136.7 Hz, 1C, C6).

^{31}P NMR (162 MHz; CD_3OD) δ 24.96 (1 P).

MS: m/z = 267.1 $[M-H]^-$, 535.3 $[2M-H]^-$.

Deprotected phosphonate (31)

Compound **49** (40 mg, 0.157 mmol) was dissolved in D_2O (0.7 mL) and the solution was transferred into an NMR tube. The reaction was followed by ^{1}H-NMR. In order to avoid the formation of NaCl, DCl was not added to catalise the deprotection and the solution was heated at 90°C for 3 days. Then the pH was neutralized adding

NaOH 5% and the solvent was liophilized to give **31** in quantitative yields as a mixture of the anomers (α/β = 1.5).

^{1}H NMR (400 MHz, D_2O) δ 5.28 (d, $J_{1,2}$ = 4.6 Hz, 1H, H1β), 5.25 (d, $J_{1,2}$ = 1.8 Hz, 1H, H1α), 4.12 – 3.96 (m, 4H, 2 H2, H4, H3), 3.91 – 3.85 (m, 1H, H3), 3.76 (dd, *J* = 13.0, 7.1 Hz, 1H, H4), 1.97 – 1.73 (m, 2H, H5), 1.67 – 1.40 (m, 2H, H6).

^{13}C NMR (101 MHz, D_2O) δ 100.60 (1C, C1α), 94.64 (1C, C1β), 83.46 (1 C, d $^{3}J_{C\text{-}P}$ = 17.0 Hz, C4α), 81.41 (1C, d $^{3}J_{C\text{-}P}$ = 18.3 Hz, C4β), 81.63, 78.88, 77.19, 76.09 (4C, 2 C2, 2 C3), 28.98, 27.62 (2C, C5), 25.16, 23.84 (2C, C6).

^{31}P NMR (162 MHz; D_2O) δ 26.53, 26.31 (2 P)

MS: m/z = 227.1 $[M\text{-}H]^-$.

α-Hydroxy-diethylphosphonate (50)

To a stirred solution of the freshly prepared aldehyde **46** (600 mg) in dry DCM (20 mL), dry TEA (1280 ml, 10.7 mmol, 5 eq) and diethylphosphite (828 ml, 6.42 mmol, 3 eq) were added. The reaction was stirred for 15 hours, then the solvent was evaporated and the resulting product was purified by flash column chromatography (from 1:1, PE:EtOAc to 100% EtOAc) giving **50** (796 mg, 89% yield in two steps) as a light yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 – 7.24 (m, 5H, H Ph), 5.90 (d, $J_{1,2}$ = 3.0 Hz, 1H, H1), 4.69 (d, *J* = 11.4 Hz, 1H, H'Bn), 4.61 (d, $J_{2,1}$ = 3.0 Hz, 1H, H2), 4.57 (d, *J* = 11.4 Hz, 1H, H''Bn), 4.51 – 4.44 (m, 1H, H4), 4.39 (s, 1H, H3), 4.23 – 4.12 (m, 4H, 2CH_2 Et), 4.11 – 4.03 (m, 1H, H5), 1.53 (s, 3H, Me), 1.36 – 1.22 (m, 9H, Me, 2CH_3 Et).

^{13}C NMR (101 MHz, $CDCl_3$) δ 137.54 (1C, C_{quat} Ph), 128.47 (1C, C Ph), 128.36 (1C, C Ph), 127.91 (1C, C Ph), 127.77 (1C, C Ph), 112.62 (1C, C_{quat}), 106.15 (1C, C1), 85.16 (1C, C4), 84.81 (1C, C2), 82.35 (1C, C3), 71.59 (1C, CH_2 Bn), 67.83 (1 C, d *J* = 161.1 Hz, C5), 63.09 (1C, d, *J* = 6.8 Hz, CH_2 Et), 62.75 (1C, d, *J* = 6.8 Hz, CH_2 Et), 26.74, 25.92 (2C, Me), 16.44 (2C, CH_3 Et).

^{31}P NMR (162 MHz; $CDCl_3$) δ 21.85 (1 P).

MS: m/z = 439.3 $[M+Na]^+$, 855.4 $[2M+Na]^+$.

Alcohol (51)

Compound **50** (375 mg, 0.900 mmol) was dissolved in EtOH (9 mL). Palladium hydroxide (200 mg) was added, the mixture was stirred under vacuum in order to remove the residual gas and then H_2 was added. The reaction was stirred at room temperature for 24 hours and then the catalyser was filtered on a syringe filter and washed with EtOH. The solvent was evaporated giving **51** as a colourless oil (290 mg, 99% yield).

^{1}H NMR (400 MHz, CD_3OD) δ 5.90 (d, *J* = 3.5 Hz, 1H, H1), 4.52 (d, *J* = 3.5 Hz, 1H, H2), 4.41 (s, 1H, H3), 4.24 – 4.13 (m, 5H, H5, 2CH_2 Et), 4.10 (s, 1H, H4), 1.50 (s, 3H, Me), 1.37 – 1.26 (m, 9H, CH_3 Et, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 112.13(1 C, C_{quat}), 106.33 (1C, C1), 87.19 (1C, C4), 86.36 (1C, C2), 74.93 (1C, C3), 66.87 (1C, d *J* = 165.8 Hz, C5), 62.60, 62.52 (2C, CH_2 Et), 25.54 (1C, Me), 24.55 (1C, Me), 15.31 (2C, CH_3 Et).

^{31}P NMR (162 MHz; $CDCl_3$) δ 24.20 (1 P).

MS: m/z = 349.2 $[M+Na]^+$, 375.3 $[2M+Na]^+$.

α-Hydroxy-phosphonate (52)

To a stirred solution of compound **51** (242 mg, 0.742 mmol) in dry ACN (7 mL), dry Py (600 ml, 7.42 mmol, 10 eq) and TMSBr (979 ml, 7.42 mmol, 10 eq) were added at room temperature. After 50 minutes the solution is cooled at 0°C and the reaction is quenched adding 5 mL of NH_4OH 33% under stirring. The ice bath is removed and the solvent is evaporated under reduced pressure. The product was purified by flash column chromatography (6:4, IPA: NH_4OH 33%) giving **52** (120 mg, 60% yield) as a white solid.

^{1}H NMR (400 MHz, CD_3OD) δ 5.78 (d, *J* = 4.2 Hz, 1H), 4.60 (bs, 1H, H2), 4.46 (d, *J* = 3.8 Hz, 1H, H3), 4.08 – 3.96 (m, 2H, H4, H5), 1.56 (s, 3H, Me), 1.34 (s, 3H, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 113.31 (1C, C_{quat}), 103.99 (1C, C1), 88.04 (1C, C2), 84.83 (1C, d, $^2J_{CP}$ = 13.8 Hz, C4), 73.55 (1C, C3), 67.97 (1C, d, $^1J_{CP}$ = 149.1 Hz, C5), 26.69, 26.14 (2C, Me).

^{31}P NMR (162 MHz; CD_3OD) δ 16.70 (1 P).

MS: m/z = 269.0 $[M-H]^-$.

Deprotected α-Hydroxy-phosphonate (34)

Compound **52** (40 mg, 0.157 mmol) was dissolved in D_2O (0.7 mL) and the solution was transferred into an NMR tube. The reaction was followed by ^{1}H-NMR. In order to avoid the formation of NaCl, DCl was not added to catalyse the deprotection and the solution was heated at 90°C for 5 days. Then the pH was neutralized adding NaOH 5% and the solvent was lyophilised to give **34** in quantitative yields, as a mixture of anomers for pyranose e furanose forms. Due to the overlapping of the signals of the 4 products, it was not possible to assign clearly every signal.

^{1}H NMR (400 MHz, D_2O) δ 5.08 (d, *J* = 4.5 Hz, 1H), 5.03 – 4.98 (m, 2H), 4.89 (s, 1H), 4.14 (t, *J* = 8.3 Hz, 2H), 4.10 – 3.93 (m, 4H), 3.92 – 3.84 (m, 2H), 3.80 (dd, *J* = 14.3, 3.1 Hz, 2H), 3.73 (t, *J* = 12.7 Hz, 2H), 3.63 (d, *J* = 4.3 Hz, 2H), 3.53 (t, *J* = 9.7 Hz, 2H), 3.42 (t, *J* = 7.3 Hz, 1H).

^{13}C NMR (101 MHz, D_2O) δ 100.17, 93.76, 93.73, 92.75, 92.63, 81.45, 75.53, 72.34, 71.23, 70.28, 70.01, 68.36, 68.26, 65.74.

^{31}P NMR (162 MHz; D_2O) δ 16.94 (1 P), 16.59 (1 P), 16.30 (1 P), 14.77 (1 P).

MS: m/z = 229.1 $[M-H]^-$.

α-O-Acetyl-diethylphosphonate (53)

To a stirred solution of the hydroxyphosphonate **50** (68 mg, 0.163 mmol, 1 eq) in EtOAc (1 mL), Py (20 ml, 0.245 mmol, 1.5 eq), acetic anhydride (23 ml, 0.245 mmol, 1.5 eq) and DMAP (2 mg, 0.016 mmol, 0.1 eq) were added. The reaction was stirred for 2 hours, then EtOH was added, the solvent was evaporated and the resulting product was purified by flash column chromatography (1:1, PE:EtOAc) giving **53** (67 mg, 90% yield) as a light yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.24 (m, 5H, H Ph), 5.93 (d, $J_{1,2}$ = 3.7 Hz, 1H, H1), 5.49 (t, *J* = 8.7 Hz, 1H, H5), 4.62 (d, $J_{2,1}$ = 3.7 Hz, 1H, H2), 4.57 (d, *J* = 2.4 Hz, 2H, CH_2 Bn), 4.52 – 4.44 (m, 1H, H4), 4.24 – 4.09 (m, 4H, 2CH_2 Et), 3.95 (d, *J* = 1.8 Hz, 1H, H3), 2.00 (s, 3H, CH_3 OAc), 1.58 (s, 3H, Me), 1.33 – 1.23 (m, 9H, CH_3 Et, Me).

^{13}C NMR (101 MHz, $CDCl_3$) δ 172.52 (1C, CO), 137.01 (1C, C_{quat} Ph), 128.45 (1C, C Ph), 128.42 (1C, C Ph), 127.94 (1C, C Ph), 127.79 (1C, C Ph), 113.06 (1C, C_{quat}), 106.08 (1C, C1), 84.48 (1C, C2), 82.21 (1C, C3), 81.90 (1C, C4), 71.46 (1C, CH_2

Bn), 66.69 (1C, d, J_{CP} = 165.7 Hz, C5), 63.23, 62.64 (2C, CH_2 Et), 26.75 (1C, Me), 26.04 (1C, Me), 20.57 (1C, CH_3 OAc), 16.35 (2C, CH_3 Et).

^{31}P NMR (162 MHz; $CDCl_3$) δ 17.86 (1 P).

MS: m/z = 459.3 $[M+H]^+$, 481.3 $[M+Na]^+$, 497.3 $[2M+K]^+$.

Alcohol (54)

Compound **53** (42 mg, 0.092 mmol) was dissolved in EtOH (1 mL). Palladium hydroxide (20 mg) was added, the mixture was stirred under vacuum in order to remove the residual gas and then H_2 was added. The reaction was stirred at room temperature for 24 hours and then the catalyser was filtered on a syringe filter and washed with EtOH. The solvent was evaporated giving **54** as colourless oil (28 mg, 98% yield).

^{1}H NMR (400 MHz, CD_3OD) δ 5.93 (d, $J_{1,2}$ = 3.5 Hz, 1H, H1), 5.51 (dd, J = 10.0, 7.1 Hz, 1H, H5), 4.51 (d, $J_{2,1}$ = 3.5 Hz, 1H), 4.26 (t, J = 9.2 Hz, 1H, H4), 4.24 – 4.10 (m, 4H, 2 CH_2 Et), 4.08 (s, 1H, H3), 2.13 (s, 3H, CH_3 OAc), 1.55 (s, 3H, Me), 1.37 – 1.26 (m, 9H, CH_3 Et, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 172.03 (1C, CO), 116.25 (1C, C_{quat}), 110.41 (1C, C1), 90.16 (1C, C2), 89.07 (1C, C4), 78.90 (1C, C3), 70.56 (1C, d, J = 168.1 Hz, C5), 67.30, 66.80 (2C, CH_2 Et), 29.38 (1C, Me), 28.41 (1C, Me), 23.06 (1C, CH_3 OAc), 19.20 (2C, CH_3 Et).

^{31}P NMR (162 MHz; CD_3OD) δ 18.77 (1 P).

MS: m/z = 369.2 $[M+H]^+$, 391.3 $[M+Na]^+$, 407.2 $[2M+K]^+$.

α-O-Acetyl-phosphonate (55)

To a stirred solution of compound **54** (58 mg, 0.157 mmol) in dry ACN (2 mL), dry Py (178 ml, 2.205 mmol, 14 eq) and TMSBr (207 ml, 1.575 mmol, 10 eq) were added at room temperature. After 40 minutes the solution is cooled at 0°C and the reaction is quenched adding 2 mL of Py and 2 mL of EtOH under stirring. The ice bath is removed and the solvent is evaporated under reduced pressure. The product was purified by flash column chromatography (1:1 $CHCl_3$: MeOH) giving **55** (39 mg, 52% yield) as a light yellow solid.

^{1}H NMR (400 MHz CD_3OD) δ 5.86 (d, $J_{1,2}$ = 3.5 Hz, 1H, H1), 5.58 – 5.50 (m, 1H, H5), 4.58 (d, $J_{2,1}$ = 3.5 Hz, 1H, H2), 4.31 (bs, 1H, H4), 4.26 (bs, 1H, H3), 2.11 (s, 1H, CH_3 OAc), 1.48 (s, 1H, Me), 1.28 (s, 1H, Me).
^{31}P NMR (162 MHz; CD_3OD) δ 13.56 (1 P).
MS: m/z = 311.2 $[M-H]^-$.

Deprotected α-O-Acetyl-phosphonate (35)

Compound **55** (30 mg, 0.157 mmol) was dissolved in D_2O (0.7 mL) and the solution was transferred into an NMR tube. The reaction was followed by ^{1}H-NMR. In order to avoid the formation of NaCl, DCl was not added to catalise the deprotection and the solution was heated at 90°C for X days. Then the pH was neutralized adding NaOH 5% and the solvent was lyophilised to give **35**. Unfortunately **35** is not stable, in particular the acetate group migrated leading to a mixture of compound; a clear NMR signal assignation was not possible.
^{31}P NMR (162 MHz; CD_3OD) δ 10.24 (bs P).

Triflate (56)

To a stirred solution of **45** (110 mg, 0.389 mmol, 1eq) in dry DCM (4 mL), DTBMP (200 mg, 0.975 mmol, 2.5 eq) was added. The solution is cooled at -75°C and triflic anhydride (132 mL, 0.785 mmol, 2 eq) was added by syringe; the temperature was increased to 0°C and after 30 minutes the reaction mixture was directly loaded on a silica gel column and quickly filtrated with (6:4, PE:EtOAc) giving **56**. The compound wasn't characterised and it was used immediately for the next step because of its instability.

Diethyldifluoromethylphosphonate (57)

A 2 M solution of LDA in THF (1.17 mL, 2.34 mmol, 6 eq) was diluted in dry THF (2 mL) and cooled at -78°C. HMPA (407 mL, 2.34 mmol, 6 eq) and diethyldifluoromethylphosphonate (367 mL, 2.34 mmol, 6 eq) were dissolved in dry THF (1 mL) and the solution was cooled at -78°C and then slowly transferred by cannula to the cold LDA solution. The mixture was stirred at -78°C for 20 minutes, then a dry THF (2 mL) solution of the freshly prepared triflate **56** was cooled at -

78°C and transferred slowly via cannula to the cold $LiCF_2PO(OEt)_2$ solution. The mixture was stirred at -78°C for 25 minutes, then it was quenched adding NH_4Cl saturated solution. The product was extracted 4 times with EtOAc, the organic phase is dried over Na_2SO_4, filtered and concentrated under reduced pressure. The product was purified by flash column chromatography (7:3, PE:EtOAc) giving **57** (50 mg, 30% yield in two steps) as a light yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.18 (m, 5H, H Ph), 5.91 (d, $J_{1,2}$ = 3.8 Hz, 1H, H1), 4.67 – 4.58 (m, 4H, H2, H4, CH_2 Bn), 4.35 – 4.20 (m, 4H, CH_2 Et), 4.03 (s, 1H, H3), 2.69 – 2.45 (m, 2H, H5), 1.52 (s, 3H, Me), 1.42 – 1.34 (m, 6H, CH_3 Et), 1.32 (s, 3H, Me).

^{13}C NMR (101 MHz, $CDCl_3$) δ 137.25 (1C, C_{quat} Ph), 128.52 (1C, C Ph), 128.46 (1C, C Ph), 128.03 (1C, C Ph), 127.98 (1C, C Ph), 127.86 (1C, C Ph), 112.49 (1C, C_{quat}), 106.05 (1C, C1), 85.49 (1C, C3), 84.64 (1C, C2), 78.68 (1C, C4), 71.82 (1C, CH_2 Bn), 64.60, 64.55 (2C, CH_2 Et), 26.75 (1C, Me), 25.95 (1C, Me), 22.92 (d, J = 27.1 Hz, 1C, C5), 16.45 (2C, CH_3 Et).

^{31}P NMR (162 MHz; $CDCl_3$) δ 6.32 (1 P, t, $J_{P\text{-}F}$ = 106.8 Hz).

^{19}F(376 MHz; $CDCl_3$) -111.77 (1 F, ddd, *J* 107.4, 25.3 and 15.4), -112.24 (1 F, ddd, *J* 106.1, 24.2 and 16.3).

MS: m/z = 451.3 $[M+H]^+$, 473.3 $[M+Na]^+$, 923.3 $[2M+Na]^+$.

Alcohol (58)

Compound **57** (40 mg, 0.218 mmol) was dissolved in EtOH (1 mL). Palladium hydroxide (40 mg) was added, the mixture was stirred under vacuum in order to remove the residual gas and then H_2 was added. The reaction was stirred at room temperature for 24 hours and then the catalyser was filtered on a syringe filter and washed with EtOH. The solvent was evaporated giving **58** as colourless oil (32 mg, 98%).

^{1}H NMR (400 MHz, CD_3OD) δ 5.88 (d, $J_{1,2}$ = 3.8 Hz, 1H, H1), 4.52 (d, $J_{2,1}$ = 3.8 Hz, 1H, H2), 4.37 – 4.24 (m, 5H, H4, $2CH_2$ Et), 4.15 (s, 1H, H3), 2.71 – 2.41 (m, 2H, H5), 1.49 (s, 3H, Me), 1.41 – 1.36 (m, 6H, $2CH_3$ Et), 1.30 (s, 3H, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 113.27 (1C, C_{quat}), 107.54 (1C, C1), 87.69 (1C, C2), 82.58 (1C, C4), 79.46 (1C, C3), 66.33, 66.26 (2C, CH_2 Et), 39.31 (1C, C5), 26.82 (1C, Me), 25.86 (1C, Me), 16.75, 16.70 (2C, CH_3 Et).

^{31}P NMR (162 MHz; CD_3OD) δ 6.41 (1 P, t, $J_{P\text{-}F}$ = 217.3 Hz).

^{19}F(376 MHz; CD_3OD) -113.26 (1 F, app. t, *J* 20.4), -113.55 (1 F, app. t, *J* 19.4).

MS: m/z = 383.2 $[M+Na]^+$, 743.2 $[2M+Na]^+$.

Difluoromethylphosphonate (59)

To a stirred solution of compound **58** (18 mg, 0.05 mmol) in dry ACN (1 mL), dry Py (130 ml, 1.6 mmol, 32 eq) and TMSBr (172 ml, 1.3 mmol, 26 eq) were added at room temperature. After 10 hours the solution is cooled at 0°C and the reaction is quenched adding 4 mL of Py/H_2O 1:1 under stirring. The ice bath is removed and the solvent is evaporated under reduced pressure. The product was purified by flash column chromatography (6:4, IPA: NH_4OH 33%) giving **59** (13 mg, 85% yield) as a white solid.

^{1}H NMR (400 MHz, CD_3OD) δ 5.85 (d, $J_{1,2}$ = 3.9 Hz, 1H, H1), 4.55 (d, $J_{2,1}$ = 3.9 Hz, 1H, H2), 4.34 (t, *J* = 6.6 Hz, 1H, H4), 4.16 (s, 1H, H3), 2.46 – 2.25 (m, 2H, H5), 1.38 (s, 3H, Me), 1.18 (s, 3H, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 119.48 (1 C, d, J_{CF} = 219.5 Hz, C6), 113.47 (1C, C_{quat}), 107.00 (1C, C1), 88.24 (1C, C2), 83.98 (1C, C4), 79.60 (1C, C3), 40.28 (1C, C5), 27.16 (1C, Me), 26.21 (1C, Me).

^{31}P NMR (162 MHz; CD_3OD) δ 4.08 (1 P, t, $J_{P\text{-}F}$ = 89.8 Hz).

^{19}F(376 MHz; CD_3OD) -113.34 (1 F, app. t, *J* 20.1), -113.59 (1 F, app. t, *J* 20.0).

MS: m/z = 303.2 $[M-H]^-$.

Deprotected difluoromethylphosphonate (36)

Compound **59** (10 mg, 0.033 mmol) was dissolved in D_2O (0.6 mL) and the solution was transferred into an NMR tube. The reaction was followed by ^{1}H-NMR. The solution at pH 4 was heated at 90°C for 1 day. Then the pH was neutralized adding NaOH 5% and the solvent was lyophilized to give **36** in quantitative yields.

^{1}H NMR (400 MHz, D_2O) δ 5.07 (d, *J* = 4.3 Hz, 1H, H1β), 5.06 (d, *J* = 1.7 Hz, 1H, H1α), 4.28 – 4.20 (m, 1H, H4α), 3.99 – 3.92 (m, 1H, H4β), 3.89 – 3.80 (m, 3H, H2α, H2β, H3β), 3.75 – 3.70 (m, 1H, H3α), 2.40 – 2.07 (m, 2H, H5).

^{13}C NMR (101 MHz, D_2O) δ 119.22 (1 C, d J_{CF} = 271.0 Hz, C6), 99.43 (1 C, C1α), 93.68 (1C, C1β), 79.43 (1C, C3β), 78.07 (1C, C3α), 76.33 (1C, C2β), 75.25 (1C, C4α), 73.96 (1C, C2α), 73.32 (1C, C4β), 36.14 (1 C, d, J_{CF} = 163.6 Hz, C5).

^{31}P NMR (162 MHz; D_2O) δ 4.26 (1 P, t, $J_{P\text{-}F}$ = 94.8 Hz), 4.19 (1 P, t, $J_{P\text{-}F}$ = 94.8 Hz).

^{19}F(376 MHz; D_2O) -112.36 - -110.81 (2F, m), -114.59 - -113.42 (2F, m).

MS: m/z = 263.1 $[M\text{-}H]^-$.

Iodide (33)

To a stirred solution of **45** (600 mg, 2.140 mmol, 1eq) in dry Toluen (20 mL), tryphenylphosphine (2.95 g, 10.7 mmol, 5 eq), imidazole (0.459 mg, 6.42 mmol, 3 eq) and then iodine (2.28 g, 8.56 mmol, 4 eq) were added. The reaction mixture was stirred for 1 hour at 100°C, then 20 mL of $NaHCO_3$ saturated solution were added. Under vigorous stirring, iodine was added until the mixture become dark coloured, in order to oxidize the trypenyl phosphine in excess. Then the iodine in excess is reduced adding a saturated solution of $Na_2S_2O_4$ (about 50 mL). The mixture was extracted 3 times with toluen; the organic phase was dried with Na_2SO_4, filtered and concentrated under reduced pressure. The product was purified by flash column chromatography (from 9.5:0.5 to 8:2, PE:EtOAc) giving **33** (784 mg, 94% yield) as a white solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.79 – 7.27 (m, 5H, H Ph), 5.98 (d, $J_{1,2}$ = 3.7 Hz, 1H, H1), 4.69 – 4.60 (m, 3H, H2, CH_2 Bn), 4.38 (t, *J* = 7.5 Hz, 1H, H4), 4.15 (s, 1H, H3), 3.40 (d, *J* = 2.7 Hz, 1H, H5'), 3.38 (s, 1H, H5''), 1.54 (s, 3H, Me), 1.33 (s, 3H, Me).

^{13}C NMR (101 MHz, $CDCl_3$) δ 137.08 (1C, C_{quat} Ph), 128.55 (1C, C Ph), 128.47 (1C, C Ph), 128.05 (1C, C Ph), 127.90 (1C, C Ph), 112.70 (1C, C_{quat}), 106.52 (1C, C1), 85.68 (1C, C4), 84.66 (1C, C2), 83.82 (1C, C3), 71.59 (1C, CH_2 Bn), 27.06 (1C, Me), 25.95 (1C, Me), 6.32 (1C, C5).

MS: m/z = 413.2 $[M+Na]^+$, 803.2 $[2M+Na]^+$.

Ammonium diethylthiophosphate

Sulfur (500 mg, 15.596 mmol, 1 eq) and NH_4HCO_3 (1.233 g, 15.596 mmol, 1 eq) were suspended in Et_2O (25 mL) and EtOAc (25 mL). Diethylphosphite (2 mL, 15.596 mmol, 1 eq) was added and then the mixture was vigorously stirred for 24 hours. The solvent was evaporated under reduced pressure, giving **$NH_4SPO(OEt)_2$** as a white solid in quantitative yields. The characterization was in agreement with that one reported in literature.[249]

Diethylthiophosphate (61)

Compound **33** (354 mg, 0.908 mmol) was dissolved in dry DMF (10 mL). Ammonium diethylthiophosphate ($NH_4SPO(OEt)_2$) (668 mg, 3.398 mmol, 4 eq) was added and the reaction was stirred at 80°C for 18 hours. Then the solvent was evaporated under reduced pressure and the product was purified by flash column chromatography (6:4, PE:EtOAc) giving **61** (219 mg, 56% yield) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 – 7.20 (m, 5H, H Ph), 5.88 (d, $J_{1,2}$ = 3.8 Hz, 1H, H1), 4.60 (d, $J_{2,1}$ = 3.8 Hz, 1H, H2), 4.57 (s, 2H, CH_2 Bn), 4.27 (t, $J_{2,1}$ = 6.6 Hz, 1H, H4), 4.19 – 4.03 (m, 4H, CH_2 OEt), 3.98 (d, J = 1.2 Hz, 1H, H3), 3.18 – 3.00 (m, 2H, H5), 1.49 (s, 3H, Me), 1.36 – 1.29 (m, 2H, Me, 2CH_3 OEt).

^{13}C NMR (101 MHz, $CDCl_3$) δ 137.22 (1C, C_{quat} Ph), 128.70 (1C, C Ph), 128.19 (1C, C Ph), 128.00 (1C, C Ph), 112.81 (1C, C_{quat}), 106.34 (1C, C1), 84.78 (1C, C2), 84.51 (1C, C4), 84.01 (1C, C3), 71.82 (1C, CH_2 Bn), 63.92 (2C, CH_2 OEt), 33.25 (d, $J_{C\text{-}P}$ = 3.4 Hz, 1C, C5), 27.11 (1C, Me), 26.15 (1C, Me), 16.25 (2C, CH_3 OEt).

^{31}P NMR (162 MHz; $CDCl_3$) δ 26.99 (1 P).

MS: m/z = 433.3 $[M+H]^+$, 455.3 $[M+Na]^+$, 887.4 $[2M+Na]^+$.

Iodide (62)

Methyl-D-arabinofuranoside (**27**/**28**) (1.03 g, 6.165 mmol) was dissolved in dry THF (60 mL). Triphenylphosphine (3.23 g, 12.33 mmol, 2 eq), imidazole (840 mg, 12.33 mmol, 2 eq) and iodine (3.13 g, 12.33 mmol, 2 eq) were added and the reaction was refluxed under stirring for 25 minutes. Then Na_2CO_3 saturated solution (50 mL) and $Na_2S_2O_3$ saturated solution (until complete discolouration) were added

under stirring. THF was evaporated under reduced pressure and the product was extracted with EtOAc; the organic phase was dried over Na_2SO_4, filtered and concentrated under reduced pressure. The product was purified by flash column chromatography (1:1, PE:EtOAc) giving **62** (1.35 g, 80% yield) as a.

^{1}H NMR (400 MHz, CD_3OD) δ 4.72 (d, $J_{1,2}$ = 1.7 Hz, 1H, H1), 3.94 (dd, $J_{2,3}$ = 3.8, $J_{2,1}$ = 1.7 Hz, 1H, H2), 3.77 – 3.67 (m, 2H, H3, H4), 3.44 – 3.37 (m, 1H, H5'), 3.31 (s, 3H, OMe), 3.28 – 3.24 (m, 1H, H5'').

^{13}C NMR (101 MHz, CD_3OD) δ 109.09 (1C, C1), 82.46 (1C, C4), 82.26 (1C, C2), 80.79 (1C, C3), 53.97 (1C, OMe), 5.11 (1C, C5).

MS: m/z = 275.5 $[M+H]^+$.

Diethylthiophosphate (63)

Compound **62** (1.35 g, 4.927 mmol) was dissolved in dry DMF (30 mL). Previously synthesised ammonium diethylthiophosphate ($NH_4SPO(OEt)_2$) (3.69 g, 19.708 mmol, 4 eq) was added and the reaction was stirred at 80°C for 2 hours. Then the solvent was evaporated under reduced pressure and the product was purified by flash column chromatography (from 1:1, PE:EtOAc to 100% EtOAc) giving **63** (950 mg, 61% yield) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.87 (s, 1H, H1), 4.25 – 3.95 (m, 7H, H2, H3, H4, 2 CH_2 Et), 3.38 (s, 3H, OMe), 3.24 – 3.05 (m, 2H, H5), 1.35 (t, J = 7.0 Hz, 9H, 2 CH_3 Et).

^{13}C NMR (101 MHz, $CDCl_3$) δ 108.75 (1C, C1), 82.76 (1C, C4), 81.02 (1C, C21), 78.96 (1C, C3), 64.38 (2C, CH_2 Et), 55.27 (1C, OMe), 29.66 (1C, C5), 15.97 (2C, CH_3 Et).

^{31}P NMR (162 MHz; $CDCl_3$) δ 28.68 (1 P).

MS: m/z = 339.2 $[M+Na]^+$, 355.1 $[M+K]^+$, 655.2 $[2M+Na]^+$.

Peracetylated diethylthiophosphate (64)

To a stirred solution of compound **63** (920 mg, 2.912 mmol) in Ac_2O (27 mL), acetic acid (7 mL) and H_2SO_4 (370 mL) were added. The solution was stirred for 30 minutes, then it was cooled at 0°C and it was neutralized adding Na_2CO_3. The mixture was extracted 5 times with EtOAc and brine; the organic phase was dried

with Na_2SO_4, filtered and concentrated under reduced pressure. The product was purified by flash column chromatography (from 6:4 to 4:6, PE:EtOAc) giving **64** (752 mg, 60% yield) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.16 (s, 1H, H1), 5.18 (d, $J_{1,2}$ = 1.4 Hz, 1H, H2), 5.04 (dd, J = 4.8, 1.1 Hz, 1H, H3), 4.43 – 4.38 (m, 1H, H4), 4.24 – 4.09 (m, 4H, CH_2 Et), 3.31 – 3.08 (m, 2H, $2CH_3$ Et), 2.13 – 2.10 (m, 9H, CH_3, OAc), 1.36 (t, J = 7.0 Hz, 9H, $2CH_3$ Et).

^{13}C NMR (101 MHz, $CDCl_3$) δ 169.89, 169.45, 169.16 (3C, C_{quat} CO), 99.14 (1C, C1), 83.15 (1C, C4), 80.76 (1C, C3), 78.56 (1C, 2), 63.88, 63.82 (2C, CH_2 Et), 32.52 (1C, C5), 21.00, 20.71, 20.68 (3C, CH_3 OAc), 16.04, 15.97 (2C, CH_3 Et).

^{31}P NMR (162 MHz; $CDCl_3$) δ 26.48 (1 P).

MS: m/z = 451.1 $[M+Na]^+$, 879.2 $[2M+Na]^+$.

Thiophosphate (65)

To a stirred solution of compound **64** (236 mg, 0.552 mmol) in dry ACN (5.5 mL), dry Py (721 ml, 8.83 mmol, 16 eq) and TMSBr (946 ml, 7.17 mmol, 13 eq) were added at room temperature. After 1 hour the solution is cooled at 0°C and the reaction is quenched adding 4 mL of Py/H_2O 1:1 under stirring. The ice bath is removed and the solvent is evaporated under reduced pressure. The product was purified by flash column chromatography (8:2, MeOH: TEA) giving **65** (200 mg, 63% yield) as a white solid.

^{1}H NMR (400 MHz, CD_3OD) δ 6.12 (s, 1H, H1), 5.14 (s, 2H, H2, H3), 4.57 – 4.51 (m, 1H, H4), 3.29 – 3.09 (m, 3H,) , 2.10 (bs, 9 H, H OAc).

^{13}C NMR (101 MHz, CD_3OD) δ 170.16, 169.78, 169.55 (3C, C_{quat} CO), 99.08 (1C, C1), 83.73 (1C, C4), 80.93 (1C, C3), 78.63 (1C, C2), 32.21 (1C, C5), 19.54, 19.36, 19.22 (3C, CH_3 OAc).

^{31}P NMR (162 MHz; CD_3OD) δ 16.73 (1 P).

MS: m/z = 371.2 $[M-H]^-$.

Deprotected thiophosphate (37)

Compound **65** (50 mg, 0.087 mmol) was dissolved in dry MeOH (1 mL). Na (5.2 mg, 0.218 mmol, 2.5 eq) was added and the reaction was stirred for 2 hours. Then

the solvent was evaporated under reduced pressure and then the product was purified by flash column chromatography (8:2, EtOH:TEA) giving **37** as triethylammonium salt. The obtained compound was purified on a 1x26 cm IR120 Na^+ form, and slowly eluted with MeOH giving compound **7** as sodium salt (15 mg, 80% yield).

^{1}H NMR (400 MHz, D_2O) δ 4.77 (d, *J* = 1.2 Hz, 1H, H1α), 4.50 (d, *J* = 7.9 Hz, 1H, H1β), 4.13 – 4.09 (m, 1H, H4α), 4.04 – 3.97 (m, 1H, H4β), 3.79 (dd, *J* = 10.0, 3.1 Hz, 1H, H2α), 3.66 (t, *J* = 8.3 Hz, 1H, H2β), 3.61 (dd, *J* = 10.0, 3.0 Hz, 1H, H3α), 3.35 (dd, *J* = 8.5, 2.9 Hz, 1H, H3β), 3.01 (dd, *J* = 14.6, 1.4 Hz, 1H, H5'α), 2.69 (dd, *J* = 14.3, 2.1 Hz, 1H, H5'β), 2.56 (dd, *J* = 14.5, 5.6 Hz, 1H, H5''β), 2.41 (dd, *J* = 15.2, 3.7 Hz, 1H, H5''α).

13C NMR (101 MHz, D_2O) δ 74.26 (1C, C1β), 73.66 (1C, C2β), 73.47 (2C, C1α, C3β), 70.88 (1C, C2α), 69.41 (1C, C3α), 69.25 (1C, C4α), 67.88 (1C, C4β), 29.36 (1C, C5β), 28.57 (1C, C5β).

^{31}P NMR (162 MHz; D_2O) δ 20.5 (1 P).

MS: m/z = 245.1 $[M-H]^-$.

Sulphate (66)

To a stirred solution of **45** (80 mg, 0.285 mmol, 1eq) in dry DMF (3 mL), SO_3Py (454 mg, 2.854 mmol, 10 eq) was added. The reaction mixture was stirred for 1 hour and then it was neutralized adding NaOH 5%. The mixture was extracted 6 times with brine and EtOAc; the organic phase is dried with Na_2SO_4, filtered and concentrated under reduced pressure. The product was purified by flash column chromatography (95:0.5, EtOAc:MeOH) giving **66** (85 mg, 90% yield) as a white solid.

^{1}H NMR (400 MHz, CD_3OD) δ 7.46 – 7.17 (m, 5H, H Ph), 5.89 (d, $J_{1,2}$ = 3.9 Hz, 1H), 4.69 (d, $J_{2,1}$ = 3.9 Hz, 1H, H2), 4.60 (s, 2H, CH_2 Bn), 4.35 (td, $J_{4,5}$ = 6.7 Hz, $J_{4,3}$ = 1.8 Hz, 1H, H4), 4.15 (d, $J_{5,4}$ = 6.7 Hz, 2H, H5), 4.09 (d, $J_{3,4}$ = 1.8 Hz, 1H, H3), 1.51 (s, 1H, Me), 1.30 (s, 1H, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 138.87 (1C, C_{quat} Ph), 129.37 (1C, C Ph), 129.00 (1C, C Ph), 128.79 (1C, C Ph), 113.70 (1C, C_{quat}), 107.32 (1C, C1), 85.77 (1C, C2),

84.67 (1C, C4), 84.13 (1C, C3), 72.58 (1C, CH_2 Bn), 68.46 (1C, C5), 27.18 (1C, Me), 26.23 (1C, Me).
MS: m/z = 359.2 $[M-H]^-$.

Alcohol (67)

Compound **66** (70 mg, 0.212 mmol) was dissolved in MeOH (4 mL). Palladium hydroxide (70 mg) was added, the mixture was stirred under vacuum in order to remove the residual gas and then H_2 was added. The reaction was stirred at room temperature for 24 hours and then the catalyser was filtered on a syringe filter and washed with MeOH. The solvent was evaporated giving **67** as a white solid (53 mg, 98%).

^{1}H NMR (400 MHz, CD_3OD) δ 5.90 (d, $J_{1,2}$ = 3.8 Hz, 1H, H1), 4.53 (d, $J_{2,1}$ = 3.8 Hz, 1H, H2), 4.22 (s, 1H, H3), 4.20 – 4.09 (m, 3H, H4, H5), 1.50 (s, 3H, Me), 1.29 (s, 3H, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 113.51 (1C, C_{quat}), 107.34 (1C, C1), 87.86 (1C, C2), 87.30 (1C, C4), 76.52 (1C, C3), 68.60 (1C, C5), 27.03 (1C, Me), 26.10 (1C, Me).

MS: m/z = 269.1 $[M-H]^-$.

Deprotected sulphate (38)

Compound **67** (40 mg, 0.157 mmol) was dissolved in D_2O (0.7 mL) and the solution was transferred into an NMR tube. The reaction was followed by ^{1}H-NMR. In order to avoid the formation of NaCl, DCl was not added to catalyse the deprotection and the solution was heated at 90°C for 5 days. Then the pH was neutralized adding NaOH 5% and the solvent was lyophilised to give **38** in quantitative yields.

^{1}H NMR (400 MHz, D_2O) δ 5.13 (bs, 1H, H1β), 5.08 (d, J = 2.4 Hz, 1H, H1α), 4.12 – 4.02 (m, 3H, H5'α, H5'β, H3a), 3.99 – 3.91 (m, 4H, H5"α, H5"β, H2β, H2β), 3.90 – 3.81 (m, 3H, H2α, H3β, H4α).

^{13}C NMR (101 MHz, D_2O) δ 100.93 (1C, C1α), 95.19 (1C, C1β), 80.97 (1C, C2α), 80.44 (1C, C3α), 78.62 (1C, C3β), 75.76 (1C, C2β), 75.27 (1C, C4α), 73.92 (1C, C4β), 68.77 (1C, C5β), 67.38 (1C, C5α).

MS: m/z = 229.1 $[M-H]^-$.

Thioacetate (68)

Compound **33** (530 mg, 1.359 mmol) was dissolved in dry DMF (4 mL). Potassium thioacetate (388 mg, 3.398 mmol, 2.5 eq) was added and the reaction was stirred at 80°C for 1 hour. Then the solvent was evaporated under reduced pressure and the product was purified by flash column chromatography (9:1, PE:EtOAc) giving **68** (450 mg, 98% yield) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 – 7.09 (m, 5H, H Ph), 5.89 (d, *J* = 3.9 Hz, 1H, H1), 4.63 (d, *J* = 3.9 Hz, 1H, H2), 4.60 – 4.52 (m, 2H, CH_2 Bn), 4.17 (td, *J* = 7.0, 2.1 Hz, 1H, H4), 3.89 (d, *J* = 2.1 Hz, 1H, H3), 3.23 (ddd, *J* = 33.3, 13.7, 7.1 Hz, 2H, H5), 2.33 (s, 3H, OAc), 1.55 (s, 3H, Me), 1.32 (s, 3H, Me).

^{13}C NMR (101 MHz, $CDCl_3$) δ 192.29 (1C, C_{quat} CO), 134.51 (1C, C_{quat} Ph), 125.94 (1C, C Ph), 125.42 (1C, C Ph), 125.28 (1C, C Ph), 110.20 (1C, C_{quat}), 103.38 (1C, C1), 82.30 (1C, C2), 81.58 (1C, C3), 81.02 (1C, C4), 69.07 (1C, CH_2 Bn), 29.38 (1C, C5), 27.96 (1C, CH_3 OAc), 24.41 (1C, Me), 23.51 (1C, Me).

MS: m/z = 361.2 $[M+Na]^+$.

Thiol (69)

Compound **68** (357 mg, 1.056 mmol) was dissolved in dry MeOH (10 mL). Na (2,5 mg, 0.106 mmol, 0.1 eq) was added and the reaction was stirred 1 hour. Then silica gel was added, the solvent was evaporated under reduced pressure and then the product was purified by flash column chromatography (8:2, PE:EtOAc) giving **69** (270 mg, 86% yield) as a yellow oil. The disulfide was formed as secondary compound (78 mg, 13% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.60 – 6.94 (m, 5H, H Ph), 5.92 (d, $J_{1,2}$ = 3.9 Hz, 1H, H1), 4.63 (d, $J_{2,1}$ = 3.9 Hz, 1H, H2), 4.58 (d, *J* = 3.9 Hz, 2H, CH_2 Bn), 4.46 – 4.38 (m, 1H, H4), 4.05 (d, *J* = 1.1 Hz, 1H, H3), 3.00 (ddd, *J* = 22.4, 13.6, 7.3 Hz, 2H, H5), 1.50 (s, 3H, Me), 1.30 (s, 3H, Me).

^{13}C NMR (101 MHz, $CDCl_3$) δ 137.17 (1C, C_{quat} Ph), 128.54 (1C, C Ph), 128.02 (1C, C Ph), 127.85 (1C, C Ph), 112.72 (1C, C_{quat}), 105.87 (1C, C1), 86.46 (1C, C4), 84.87 (1C, C2), 83.55 (1C, C3), 71.61 (1C, CH_2 Bn), 27.28 (1C, C5), 27.07 (1C, Me), 26.14 (1C, Me).

MS: m/z = 319.2 $[M+Na]^+$.

Sulphonate (70)

Compound **69** (128 mg, 0.432 mmol) was suspended in $NaHCO_3$ saturated solution (1 mL). A 0.1 M TEMPO solution in ACN (864 mL, 0.086 mmol, 0.2 eq) and a 0.5 M KBr solution (173 mL, 0.086 mmol, 0.2 eq) were added; then NaOCl 0.35 M solution (2.5 mL, 0.864 mmol, 2 eq) was added and the reaction was stirred for 10 minutes. Then EtOH was added and the solvent was evaporated under reduced pressure; the product was purified by flash column chromatography (from 100% EtOAc to 8:2, EtOAc:EtOH) giving **70** (135 mg, 85% yield) as a light yellow solid.

^{1}H NMR (400 MHz, CD_3OD) δ 7.81 – 6.63 (m, 5H, H Ph), 5.89 (d, $J_{1,2}$ = 3.8 Hz, 1H, H1), 4.69 – 4.57 (m, 4H, H2, H4, CH_2 Bn), 4.34 (s, 1H, H3), 3.35 (dd, $J_{5',5''}$ = 13.8, $J_{5',4}$ = 8.9 Hz, 1H, H5'), 3.17 (dd, $J_{5'',5'}$ = 13.8, $J_{5'',4}$ = 4.6 Hz, 1H, H5''), 1.49 (s, 3H, Me), 1.29 (s, 3H, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 137.82 (1C, C_{quat} Ph), 127.96 (1C, C Ph), 127.59 (1C, C Ph), 127.37 (1C, C Ph), 111.92 (1C, C_{quat}), 106.11 (1C, C1), 84.68 (1C, C3), 84.30 (1C, C2), 81.92 (1C, C4), 71.25 (1C, CH_2 Bn), 54.65 (1C, C5), 25.49 (1C, Me), 24.50 (1C, Me).

MS: m/z = 343.3 $[M-H]^-$.

Alcohol (71)

Compound **70** (88 mg, 0.280 mmol) was dissolved in EtOH (3 mL). Palladium hydroxide (80 mg) was added, the mixture was stirred under vacuum in order to remove the residual gas and then H_2 was added. The reaction was stirred at room temperature for 24 hours and then the catalyser was filtered on a syringe filter and washed with EtOH. The solvent was evaporated giving **71** as a white solid (64 mg, quantitative yields).

^{1}H NMR (400 MHz, CD_3OD) δ 5.89 (d, $J_{1,2}$ = 3.8 Hz, 1H, H1), 4.53 (d, $J_{2,1}$ = 3.8 Hz, 1H, H2), 4.46 – 4.39 (m, 2H, H3, H4), 3.34 (dd, $J_{5',5''}$ = 14.0, $J_{5',4}$ = 5.7 Hz, 1H, H5'), 3.14 (dd, $J_{5'',5'}$ = 14.0, $J_{5'',4}$ = 4.9 Hz, 1H, H5''), 1.50 (s, 3H, Me), 1.30 (s, 3H, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 111.89 (1C, C_{quat}), 105.95 (1C, C1), 86.36 (1C, C2), 84.37 (1C, C4), 77.23 (1C, C3), 54.64 (1C, C5), 25.48 (1C, Me), 24.51 (1C, Me).

MS: m/z = 253.1 $[M-H]^-$.

Deprotected sulphonate (39)

Compound **71** (40 mg, 0.157 mmol) was dissolved in D_2O (0.7 mL) and the solution was transferred into an NMR tube. The reaction was followed by ^{1}H-NMR. The pH 4 solution was heated at 90°C for 1 day. Then the pH was neutralized adding NaOH 5% and the solvent was liophilized to give **39** in quantitative yields.

^{1}H NMR (400 MHz, D_2O) δ 5.10 (d, J = 3.2 Hz, 1H, H1β), 5.08 (d, $J_{1,2}$ = 2.2 Hz, 1H, H1α), 4.22 (ddd, J = 7.9, 5.7, 4.6 Hz, 1H, H4α), 3.96 – 3.88 (m, 3H, H4β, H3β, H2β,), 3.85 (dd, J = 4.0, $J_{2,1}$ = 2.2 Hz, 1H, H2α), H2, 3.83 – 3.77 (m, 1H,H3α), 3.15 – 2.97 (m, 2H,H5).

^{13}C NMR (101 MHz, D_2O) δ 101.09 (1C, C1α), 95.33 (1C, C1β), 80.78 (1C, C2α), 79.08 (1C, C3α), 78.80 (1C, C4α), 77.44, 76.79, 75.34 (3C, C2β, C3β, C4β), 55.27 (1C, C5β), 53.79 (1C, C5α).

MS: m/z = 213.1 $[M-H]^-$.

Carboxylate (72)

Compound **45** (133 mg, 0.475 mmol) was suspended in $NaHCO_3$ saturated solution (1 mL). A 0.1 M TEMPO solution in ACN (950 mL, 0.095 mmol, 0.2 eq) and a 0.5 M KBr solution (190 mL, 0.095 mmol, 0.2 eq) were added; then NaOCl 0.35 M solution (2.7 mL, 0.949 mmol, 2 eq) was added and the reaction was stirred for 40 minutes. Then EtOH was added and the solvent was evaporated under reduced pressure; the product was purified by flash column chromatography (100% EtOAc) giving **72** (143 mg, 95% yield) as a light yellow solid.

^{1}H NMR (400 MHz, CD_3OD) δ 7.39 – 7.24 (m, 5H, H Ph), 5.97 (d, $J_{1,2}$ = 3.6 Hz, 1H, H1), 4.68 (d, $J_{2,1}$ = 3.6 Hz, 1H, H2), 4.65 – 4.61 (m, 3H, H4, CH_2 Bn), 4.48 (s, 1H, H3), 1.45 (s, 3H, Me), 1.27 (s, 3H, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 177.49 (1 C, C5), 141.38 (1 C, C_{quat} Ph), 132.03 (1 C, C Ph), 131.59 (1 C, C Ph), 131.51 (1 C, C Ph), 116.63 (1 C, C_{quat}), 110.54 (1 C, C1), 88.48 (1 C, C3), 87.44 (1 C, C4), 86.62 (1 C, C2), 75.20 (1C, CH_2 Bn), 28.60 (1C, Me), 28.45 (1C, Me).

MS: m/z = 293.2 $[M-H]^-$.

Alcohol (73)

Compound **72** (125 mg, 0.395 mmol) was dissolved in MeOH (5 mL). Palladium hydroxide (100 mg) was added, the mixture was stirred under vacuum in order to remove the residual gas and then H_2 was added. The reaction was stirred at room temperature for 24 hours and then the catalyser was filtered on a syringe filter and washed with MeOH. The solvent was evaporated giving **73** as a white solid (88 mg, quantitative yields).

^{1}H NMR (400 MHz, D_2O) δ 5.89 (d, $J_{1,2}$ = 3.6 Hz, 1H, H1), 4.53 (d, J = 0.5 Hz, 1H, H4), 4.51 (d, $J_{2,1}$ = 3.6 Hz, 1H, H2), 4.32 (s, 1H, H3), 1.30 (s, 1H, Me), 1.16 (s, 1H, Me).

^{13}C NMR (101 MHz, D_2O) δ 174.79 (1 C, C5), 113.15 (1 C, C_{quat}), 105.74 (1 C, C1), 85.87 (1C, C4), 83.97(1C, C2), 76.42 (1C, C3), 24.71 (1C, Me), 24.39 (1C, Me).

MS: m/z = 203.2 $[M-H]^-$.

Deprotected carboxylate (40)

Compound **73** (40 mg, 0.157 mmol) was dissolved in D_2O (0.7 mL) and the solution was transferred into an NMR tube. The reaction was followed by ^{1}H-NMR. The pH4 solution was heated at 90°C overnight. Then the pH was neutralized adding NaOH 5% and the solvent was liophilized to give **40** in quantitative yields.

^{1}H NMR (400 MHz, D_2O) δ 5.38 (d, J = 4.3 Hz, 1H, H1β), 5.34 (d, J = 1.3 Hz, 1H, H1α), 4.48 (d, J = 4.2 Hz, 1H, H4), 4.28 (t, J = 5.8 Hz, 1H, H3β), 4.23 – 4.17 (m, 2H, H3α, H4β), 4.03 (dd, J = 5.8, 4.3 Hz, 1H, H2β), 4.00 (dd, J = 2.7, 1.8 Hz, 2H, H2α).

^{13}C NMR (101 MHz, D_2O) δ 176.25 (1C, C5), 102.08 (1C, C1α), 96.55 (1C, C1β), 82.91 (1C, C4α), 80.12 (1C, C4β), 80.04 (1C, C2α), 78.82 (1C, C3α), 77.30 (1C, C3β), 75.49 (1C, C2β).

MS: m/z = 163.1 $[M-H]^-$.

Dibenzylmalonate (74)

To an -80°C solution of LiHMDS (2040 mL, 2.037 mmol, 3 eq) in dry THF (3.6 mL), dibenzylmalonate (509 ml, 2.037 mmol, 3 eq) was added and then the mixture

was stirred at for 30 minutes. The freshly prepared triflate **56** (239 mg, 0.853 mmol, 1eq) was dissolved in dry THF (3.6 mL) and slowly added by syringe to the cold dibenzylmalonate solution. The temperature was slowly increased to 0°C in 1 hour, then the reaction was stirred at 0°C for other 20 minutes, and finally at room temperature for 1.5 hours. The reaction was quenched adding NH_4Cl saturated solution and extracted with EtOAc; the organic phase is dried with Na_2SO_4, filtered and concentrated under reduced pressure. The product was purified by flash column chromatography (8.5:1.5, PE:EtOAc) giving **74** (183 mg, 50% yield in two steps) as a light yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40 – 7.24 (m, 15H, H Ph), 5.87 (d, $J_{1,2}$ = 3.9 Hz, 1H, H1), 5.21 (d, J = 12.3 Hz, 1H, H Bn), 5.16 – 5.04 (m, 3H, H Bn), 4.61 (d, $J_{2,1}$ = 3.9 Hz, 1H, H2), 4.57 – 4.47 (m, 2H, H Bn), 4.21 – 4.13 (m, 1H, H5), 3.88 – 3.76 (m, 2H, H3, H6), 2.05 (s, 1H), 1.43 (s, 1H), 1.30 (s, 1H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 169.22 (1C, C_{quat} CO), 168.85 (1C, C_{quat} CO), 137.15 (1C, C_{quat} Ph), 135.42 (1C, C_{quat} Ph), 135.34 (1C, C_{quat} Ph), 128.65 (1C, C Ph), 128.62 (1C, C Ph), 128.40 (1C, C Ph), 128.39 (1C, C Ph), 128.29 (1C, C Ph), 128.25 (1C, C Ph), 128.06 (1C, C Ph), 127.86 (1C, C Ph), 127.73 (1C, C Ph), 127.08 (1C, C Ph), 112.70 (1C, C_{quat}), 105.95 (1C, C1), 85.53 (1C, C3), 84.85 (1C, C2), 82.10 (1C, C4), 71.64 (1C, CH_2 Bn), 67.33 (2C, CH_2 Bn), 48.74 (1C, C6), 32.98 (1C, C5), 26.74 (1C, Me), 26.18 (1C, Me).

MS: m/z = 569.4 $[M+Na]^+$.

Malonate (75)

Compound **74** (119 mg, 0.218 mmol) was dissolved in EtOH (4 mL). Palladium hydroxide (100 mg) was added, the mixture was stirred under vacuum in order to remove the residual gas and then H_2 was added. The reaction was stirred at room temperature for 24 hours and then the catalyser was filtered on a syringe filter and washed with MeOH. The solvent was evaporated giving **75** as a white solid (59 mg, quantitative yields).

^{1}H NMR (400 MHz, CD_3OD) δ 5.88 (d, $J_{1,2}$ = 3.7 Hz, 1H, H1), 4.50 (d, $J_{2,1}$ = 3.7 Hz, 1H, H2), 4.07 – 3.96 (m, 2H, H3, H4), 3.59 (dd, $J_{6,5}$ = 10.4, 4.3 Hz, 1H, H6),

2.38 – 2.26 (m, 1H, H5'), 2.20 – 2.07 (m, 1H, H5''), 1.49 (s, 3H, Me), 1.29 (s, 3H, Me).

^{13}C NMR (101 MHz, CD_3OD) δ 173.18 (1C, C_{quat} CO), 172.76 (1C, C_{quat} CO), 113.26 (1C, C_{quat}), 107.41 (1C, C1), 88.10 (1C, C2), 86.98 (1C, C4), 79.31 (1C, C3), 49.66 (1C, C6) 33.95 (1C, C5), 26.74 (1C, Me), 26.00 (1C, Me).

MS: m/z = 275.2 $[M-H]^-$.

Deprotected malonate (41)

Compound **75** (40 mg, 0.157 mmol) was dissolved in D_2O (0.7 mL) and the solution was transferred into an NMR tube. The reaction was followed by ^{1}H-NMR. In order to avoid the formation of NaCl, DCl was not added to catalyse the deprotection and the solution was heated at 90°C for 5 days. Then the pH was neutralized adding NaOH 5% and the solvent was lyophilised to give **41**. The product was not stable, thus a perfect assignation of the NMR signals was not possible, due to signals overlapping of different compounds. Mass analysis suggested the occurring of intramolecular esterification (m/z = 217.1 $[M-H_2O-H]^-$).

^{1}H NMR (400 MHz, D_2O) δ 5.04 (d, *J* = 4.6 Hz, 1Hβ), 5.01 (d, *J* = 2.5 Hz, 1Hα), 3.86 – 3.74 (m, 3H, H2α, H2β, H6), 3.64 (dd, *J* = 6.2, 4.6 Hz, 1H, H3), 3.58 (dd, *J* = 8.2, 1.2 Hz, 1H), 3.52 – 3.47 (m, 1H, H4), 3.30 – 3.26 (m, 1H), 2.00 – 1.91 (m, 1H, H5'), 1.85 – 1.75 (m, 1H, H5'').

^{13}C NMR (101 MHz, D_2O) δ 100.61, 94.84, 81.63, 81.25, 80.42, 79.56, 79.49, 78.00, 76.57, 76.10, 73.59, 73.54, 72.05, 68.41, 35.56, 35.52, 33.78.

MS: m/z = 235.1 $[M-H]^-$, 257.1 $[M-2H+Na]^-$.

8.2. Akt

General methods

Reactions were carried out using the commercially available starting materials and solvents without further purification. All solvents were dried over molecular sieves, for at least 24 h prior to use. When dry conditions were required, the reaction was performed under Ar atmosphere. Thin-layer chromatography (TLC) was performed on Silica Gel 60 F254 plates (Merck) charring with a solution containing concd H_2SO_4/EtOH/H_2O in a ratio of 5:45:45 or with an oxidant mixture composed of $(NH_4)Mo_7O_{24}$ (21 g), $Ce(SO_4)_2$ (1 g), concdH_2SO_4 (31 mL) in water (500 mL). Flash column chromatography was performed on silica gel 230–400 mesh (Merck). NMR spectra were recorded at 400 MHz (1H) 100.57 MHz (13C) and 162.01 MHz (31P) on a Varian Mercury instrument.

Biological assay *were performed by Prof. A. Zaza and F. Granucci, Biotecnology department, University of Milano-Bicocca.*

Molecular modelling *was performed by Prof L. De Gioia, Biotecnology department, University of Milano-Bicocca.*

(i) The X-ray structure of the complex between Akt PH domain and PtdIns(3, 4, 5)P3 has been obtained from the Protein Data Bank (PDB code 1H10). (ii) All molecular modelling operations (structure visualization, processing, editing and analysis) have been carried out with the chemistry package Molecular Operating Environment (MOE).[250] Missing hydrogen atoms have been added and their orientation has been optimized by Molecular Mechanics (MM) calculations (while keeping heavy atoms fixed), using the MMFF94s[251] force field. The MMFF94s force field has been already shown to be well suited to study protein–ligand interactions when ligands are organic derivatives of carbohydrates.[252] (iii) An extensive systematic search on the torsion phase space of the ligands has been carried out to locate the global energy minimum for each ligand. (iv) The potential energy surface describing protein–ligand interaction has been sampled using Autodock, according to the previous studies.[253] The best 40 poses of the ligand in the active site have been considered for the refinement steps. (v) Each of the 40 protein–ligand complexes obtained at point 4 has been used as an input of a MM optimization using the MMFF94s forcefield. The ligand and all protein atoms within

10 A from the ligand has been allowed to change their position during optimization. The remainder of protein atoms has been held fixed. (vi) The protein– ligand configurations obtained after MM optimization have been used as an input of simulated annealing (SA) simulations, carried out using two subsequent temperature ramps, in order to avoid freezing of the system in local metastable minima. In the first temperature ramp the molecular system has been gradually (50 ps) warmed up to 800 K. Then, the system was subjected to 1000 ps of classical Molecular Dynamics according to the NVT ensemble. Finally, the system has been slowly brought down to 5 K (500 ps) and anMMenergy minimization has been started from the last trajectory frame. Further SA cycles (up to 5) were carried out on each system to confirm that the adopted procedure led to exhaustive potential energy function sampling. Edocking related to the formation of the protein–ligand complex has been calculated according to the following equation: Edocking 1/4 EPL _ EP _ EL being the terms on the right-hand side representative of minima on the potential energy surface (PES) of the protein–ligand system (PL), the sole protein (P) and the ligand (L).

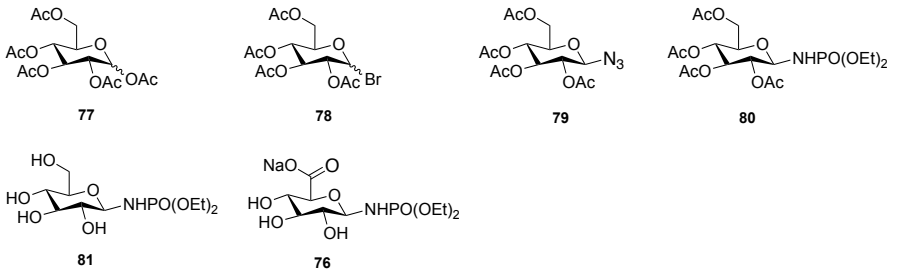


Fig. 54. Summary of the synthesised compounds.

1,2,3,4,6-Penta-O-acetyl-D-glucopyranose (77)

To a stirred solution of the **D-Glucose** (1 g, 5.552 mmol, 1 eq) in Py (1 mL), Py (10 ml), acetic anhydride (4 ml, 41.63 mmol, 1.5 eq) and DMAP (catalytic amount) were added. The reaction was stirred for 2 hours, then EtOH is added, the solvent was evaporated and the resulting product was purified by crystallization in EtOH

giving **77** (1.834 g, 85% yield) as a white solid. NMR and mass data are in agreement with those reported in literature.[150]

2,3,4,6-Tetra-O-acetyl-β-D-glucopyranosyl bromide (78)

Compound **77** (500 mg, 1.285 mmol) was dissolved in DCM (1.5 mL) and cooled at 0°C. HBr 33% solution in AcOH (0.9 mL, 5.14 mmol, 4 eq) was added, then the mixture was stirred at room temperature for 2 hours. The reaction was quenched adding ice and $NaHCO_3$ till pH neutralization. The glucopyranosyl bromide was extracted in DCM, the organic layers were collected, dryed over Na_2SO_4 anhydrous and the solvent was evaporated at reduced pressure. The obtained oil was used directly for the next step, without any further purification.

2,3,4,6-Tetra-O-acetyl-β-D-glucopyranosyl azide (79)

To a stirred solution of the freshly prepared compound **78** in DCM (5 mL) and $NaHCO_3$ saturated solution (5 mL), TBAHS (436 mg 1.285 mmol, 1 eq) and NaN_3 (250 mg, 3.855 mmol, 3 eq) were added at room temperature. After 2 hours the mixture is extracted with DCM, the organic layers were collected, dried over Na_2SO_4 anhydrous and the solvent was evaporated at reduced pressure. The product was purified by flash column chromatography (7.5:2.5 EP: EtOAc) giving **79** (289 mg, 60% yield over two steps) as a light yellow solid. NMR and mass data are in agreement with those reported in literature.[150]

2,3,4,6-Tetra-O-acetyl-β-D-glucopyranosyl diethylphosphoramidate (80)

To a solution of 2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl azide (**79**) (3.79 g, 10.15 mmol, 1.0 equiv) in dry DCM (37 mL) triethylphosphite (3.53 mL, 20.3 mmol, 2.0 equiv) was added dropwise. The reaction mixture was allowed to stir overnight at rt, then concentrated to dryness. The crude was first washed with petroleum ether, then purified by recrystallization from 4:1 EtOAc–petroleum ether affording the desired per-O-acetylated phosphoramidate (**80**) as a white solid (4.8 g, yield 98%). NMR and mass data are in agreement with those reported in literature.[150]

Glucopyranosyl diethylphosphoramidate (81)

To a solution of the per-O-acetylated phosphoramidates (**80**) (3.1 g, 6.413 mmol, 1.0 equiv) in dry methanol (15 mL) catalytic amount of metallic sodium was added. The reaction mixture was allowed to stir till completion (3 hours), then Amberlite IR 120 H^+ was added till solution neutrality. The resin was filtered off and the solvent was evaporated to dryness affording phosphoramidates **81** as amorphous white solid, (1.94 g, yield 96%). NMR and mass data are in agreement with those reported in literature.[150]

β-D-Glucuronyl diethyl phosphoramidate (76)

To a solution of glucosyl phosphoramidate (**81**) (500 mg, 1.586 mmol, 1.0 eq) in satd. aq. $NaHCO_3$, 2,2,6,6-tetramethyl-1-piperidinyloxyl (TEMPO, 49 mg, 0.317 mmol, 0.2 eq) in CH_3CN (0.1 M) and KBr (38 mg, 0.317 mmol, 0.2 eq) in satd aq $NaHCO_3$ (0.5 M, 3 mL) were added; then a solution of 0.35 M aq NaOCl (1.25 eq) was added dropwise. The reaction mixture was allowed to stir at rt. After 2 h additional 0.35 M aq NaOCl (1.25 equiv) was added dropwiseand the reaction mixture allowed to stir for other 2 hours. The reaction mixture was made acidic with 5% HCl in EtOH, then the solvent evaporated to dryness. The residue was dissolved in ethanol, the precipitate filtered off, and the filtrate was evaporated under reduced pressure. The crude was finally purified by flash chromatography with 4:5:1 EtOAc:IPA:H_2O. Hygroscopic amorphous white solid, yield (301 mg, 54%). NMR and mass data are in agreement with those reported in literature.[150]

8.3. CI-M6PR

These experiments were performed at Prof. Ben G. Davis group, Chemistry department, University of Oxford, England.

General methods

Thin layer chromatography was carried out on Merk Kieselgel 60F254 pre-coated alumium-backed plates. Visualisation of the plates was achieved using an ultraviolet lamp (λ_{max} = 254 nm) and phosphomolybdic acid/ceric sulphate (2.3%/1% in 1M H_2SO_4), potassium permanganate (0.5% in 1M NaOH), sulphuric acid (5% in ethanol) or ninhydrin (0.2% in ethanol). Flash column chromatography was carried out using Sorbsil C60 40/60 silica. "Petrol" refers to the fraction of petroleum ether boiling in the range 40- 60 ºC. "Brine" refers to a saturated aqueous solution of sodium chloride. Proton and carbon nuclear magnetic resonance spectra were recorded on a Bruker DPX 400 (400 MHz for 1H, 100.6 MHz for 13C), Bruker AV400 (400 MHz for 1H, 100.6 MHz for 13C), Bruker AVII 500 (500 MHz for 1H, 125.7 MHz for 13C). All the chemical shifts are quoted on the δ-scale in ppm using residual solvent as the internal standard. Spectra were assigned using COSY and HMQC. Low resolution mass spectra were recorded on a Micromass Platform spectrometer using electrospray ionisation (ES). m/z are reported in Daltons. High resolution mass spectra were recorded on a Walters 2790-Micromass LCT electrospray ionisation mass spectrometer using chemical ionisation techniques as stated.

Reaction with proteins were analysed in positive ion mode on a Micromass LCT mass spectrometer interfaced with a Waters 2790 Alliance HT separations module using a Phenomenex Jupiter 5u C4 300A 250 x 4.6 mm column. The sample (1 mg mL-1, 30 µL) was injected and eluted at 1.0 mL min-1 using a gradient from Solvent A (100 % water, 0.1 % formic acid) to Solvent B (100% acetonitrile, 0.1 % formic acid). The gradient was programmed as follows: 90% A (5 min isocratic) to 95% B after 15 min then isocratic for 5 min. Then to 90% A in 5 min (followed by 5 min isocratic). All solvents were degassed by sonication for 15 minutes prior to use. The eluent was split 2:8 mass spectrometer: waste. The following MS parameters were used: capillary voltage, 3000 V; sample cone voltage, 35 V; desolvation temperature, 250 oC; extraction cone, 4 V; RF lens, 250; source temperature 80 oC;

desolvation flow (N2), 540 Lh-1; no cone flow; RF DC Offset 1, 3; RF DC Offset 2, 6; volume of buffer A21 (low imidazol concentration) and 5 column volume of buffer B22 (high imidazol concentration). The eluted buffer B was concentrated on a vivaspin membrane concentrator (10 kDa molecular weight cut off) and washed with sodium phosphate buffer (50 mM, pH 7.0, 3 x 200 µL). Finally, the solution was concentrated to 100 µL and the product was characterized by LC-MS and SDS-PAGE gel.

Np276 Mutagenesis experiments.

As an example I report here the mutagenesis to obtain the 20Å ruler; for the others I followed the same procedure.

A pET28d plasmid carrying the gene for Np276 (M21I K25/94/123A N141M T201K) was modified by site-directed mutagenesis in order to introduce the mutations M61I and D41M into the proteins primary sequence using the following primers:

D41M

Forward: 5' CGAGACTTTAGTATCGTTATGTTGAGGGGTGCAGTCTTG 3'

Reverse: 5' CAAGACTGCACCCCTCAACATAACGATACTAAAGTCTCG 3‘

M61I

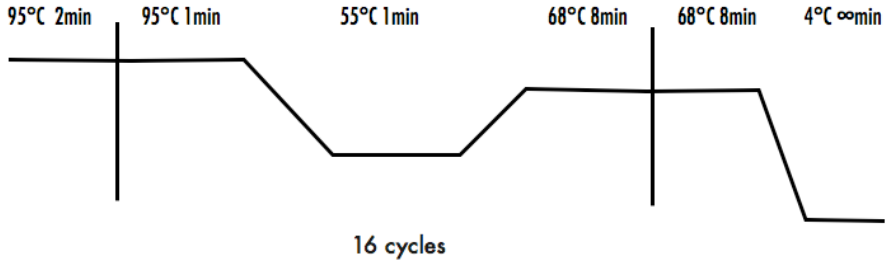
Forward: 5’-CATGGGGCGATTCTAGATGAAGC-3’

Reverse: 5’-GCTTCATCTAGAATCGCCCCATG-3’

QuikChange® II XL site-directed mutagenesis kit has been used.

To 1 µL of a 50 ng/µL DNA solution are added 5 µL of buffer, 125 ng of primers (5 µL forward and 5,25 µL revers), 1 µL dNTP solution, 3 µL of DMSO, 1 µL of PfU and milliQ water until 50 µL (29,75 µL).

The reaction run according to the following thermal cycles:



After the thermal cycles, 1 µL of DPN I solution is added, the mixture is centrifuged for 1 min and then it's incubated at 37°C for 1h.

12 µL are used for a DNA gel (+3 µL of blue dye); gel preparation: 500 mg of agarose in ~52 mL of TEA buffer, warm in microwave, then 5 µL of EtBr.

If the DNA band is visible, the DNA solution is used for the transformation in XL10 gold cells.

Transformation into XL10-Gold® Ultracompetent Cells.

1 µL of DNA solution is accurately mixed with 50 µL of XL10-Gold® Ultracompetent Cells in a pre-chili 14-ml BD Falcon polypropylene round-bottom tubes on ice.

The solution is kept at 0°C for 30 min, then for 30 sec at 42°C and then immediately at 0°C for 2 min.

In a laminar fumehood 500 µL of warm (42°C) SOC media are added and then incubated for 1h at 37°C.

After the incubation, in a sterile environment, the cells are plated on three LB/Kanamicine plates with different dilution and incubated o.n. at 37°C.

The day after, one colony is chosen and it's diluted in 5 ml LB media + 5 µL of a 50 ng/µL kanamicine solution. The falcon tube is incubated o.n. at 37°C.

DNA extraction using QIAprep spin miniprep kit.

The o.n. culture is centrifuged and the pelleted bacteria are resuspended in 250 µL of buffer P1 and transferred into a microcentrifuge tube.

250 µL of buffer P2 are added and the tubes are gently mixed 4-6 times: the cell suspension will turn blue.

350 µL of buffer N3 are added and the tubes are gently mixed 4-6 times: the solution will became cloudy.
Centrifuge for 10 min at 13000 rpm and apply the supernatant to the QIprep spin column.
Centrifuge for 60 s, discard the flow-through.
Wash the column adding 0,75 mL of buffer PE and centrifuging for 60 s. Discharge the flow through and centrifuge for other 60 s to remove the residual washing buffer.
Place the column in a clean 1.5 mL centrifuge tube and elute the DNA with 30 µL of milliQ water; add water, wait 1 min, then centrifuge for 1 min.
DNA sequencing confirmed the presence of both mutations (GeneService, Oxford):

Forward DNA sequence - T7F
atgggcagcagccatcatcatcatcacagcagcggcctggtgccgcgcggcagccat
attgacgtagaagcgctcaggcaactatatgccgcaggagagcgagactttagtatcgtt
atgttgaggggtgcagtcttggaaaacatcaatctcagtggtgcaattctacatggggcg
attctagatgaagcaaatttgcaacaggcaaatctcagtcgggctgacttaagtggggct
acgctcaatggtgcagatttaagaggggctaatttaagcgcggccgatttgagtgatgca
attcttgacaatgcaatattagaaggtgcaattcttgatgaagccgttttaaatcaggct
aatctcgcggctgctaacttggagcaggcgattcttagtcacgctaacatccgtgaagct
atgttgagtgaagctaatttggaagcagcagatttgagcggggcagatttagcgatcgcg
gatttgcatcaggcaaatctgcaccaagctgcattagaaagagccaatcttacaggagct
aatctagaagatgccaatttagaggggactattttagagggcggcaacaacaatcttgca
aaa

Transformation into B834D(E3) ultra competent cells.
1 µL of the above mutant was transformed into E. coli B834(DE3) methionineauxotrophic cells. A glycerol stock of the same cells was used to inoculate a small starter culture of LB media (25 ml) containing kanamycin (50 µg/mL). This starter culture was grown overnight at 37°C and then used to inoculate SelenoMet media (1L) containing kanamycin (50 µg/mL) and filter-sterilized L-Methionine (40 µg/mL). The cells were incubated at 37°C and grown to an optical

density value between 0.6 and 0.8, before being harvested by centrifugation (10 min, 8000 rpm, 4°C).

The cells were washed thrice by re-suspending in SelenoMet media containing kanamycin (50 µg/mL) but absent of either L-Methionine or azidohomoalanine. The cells were then transferred to pre-warmed SelenoMet media (1L) containing kanamycin (50 µg/mL) and supplemented with filter-sterilized azidohomoalanine (100µg/mL). The culture was incubated for 30min at 37°C followed by 30min at 30°C. Expression was induced with IPTG (0.24 g, 1mmol) and incubation continued overnight at 30°C (220 rpm).

The cells were then harvested by centrifugation (10 min, 8000 rpm, 4°C) and the pellets collected and incubated on ice for 1 hour with stirring following the addition of DNase (10mg), lysozyme (10mg) and protease inhibitor. The cells were then sonicated (20% power, 5 bursts of 30 s with 1 minute waiting times between bursts) and the resulting cell debris pelleted by centrifugation (30 minutes, 20,000 rpm, 4°C). Purification of the protein was achieved by incubating the cell lysate overnight with HisBind resin (10 mL) at 4 °C. The resin was washed with 150 mL of binding buffer (20 mmol Tris-HCl, 15 mmol imidazole, 0.3 M NaCl, pH 7.8) and finally eluted with 60 mL of elution buffer (20 mmol Tris-HCl, 500 mmol imidazole, 0.3 M NaCl, pH 7.8). All fractions were analyzed by SDS-PAGE (4-12% Bis-Tris gel, MES buffer).

Extra purification step was required using Q anion exchange column. Buffer A: Tridma HCl 20 mM (3.16 g/L), NaCl 150 mM (8.77 g/L), pH=8; Buffer B: Tridma HCl 20 mM (3.16 g/L), NaCl 1 M (58.4 g/L), pH=8. The fractions containing protein were combined and dialyzed twice into 4 L of phosphate buffer (50 mmol Na_2HPO_4, pH 8) to remove the imidazole. The resulting protein, now in phosphate buffer was analyzed by LCT-MS and stored at -20°C .

Sequence of NP276 M21/61I K25/94/123A D41M N141M T201K (20 Å ruler).

M G S S H H H H H H S S G L V P R G S H

I D V E A L R Q L Y A A G E R D F S I V

M L R G A V L E N I N L S G A I L H G A

I L D E A N L Q Q A N L S R A D L S G A

T L N G A D L R G A N L S A A D L S D A
I L D N A I L E G A I L D E A V L N Q A
N L A A A N L E Q A I L S H A N I R E A
M L S E A N L E A A D L S G A D L A I A
D L H Q A N L H Q A A L E R A N L T G A
N L E D A N L E G T I L E G G N N N L A
K

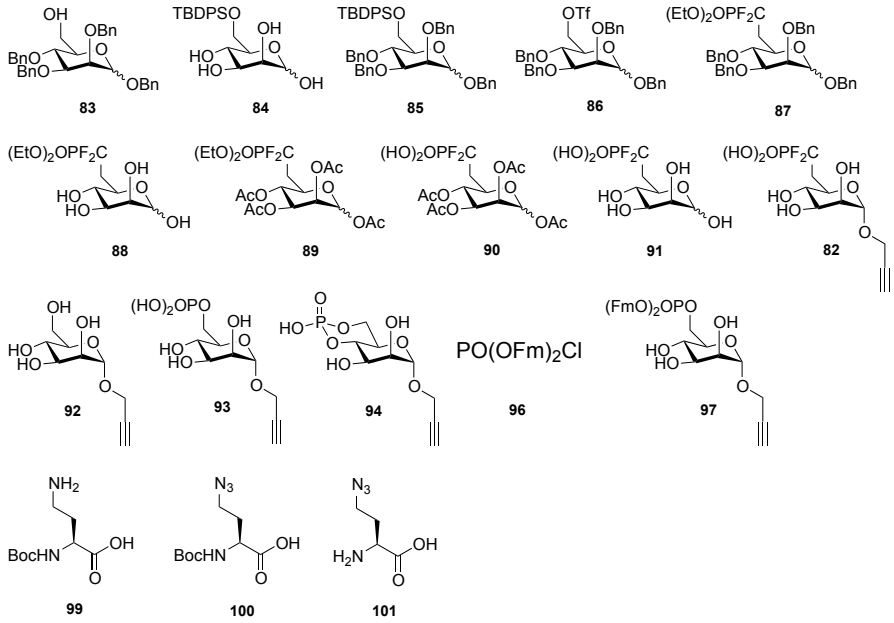


Fig. 55. Summary of the synthesised compounds.

6-O-t-Butyldiphenylsilyl-D-mannopyranose (84)

5,55 mmol of D-Mannose were dissolved in dry Py. At 0°C, N_2 atmosphere and under stirring, 0,1 eq of DMAP is added, then 1,1 eq of TBDPSCl is dropped slowly. The reaction was stirred overnight at 4°C and then quenched with ethanol. The solvent was evaporated and the product purified by Biotage 100g snap column, 100% AcOEt to obtain **84** (90%), as a white sirup.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.68-7.69 (4 H, H Ph), 7.35-7.36 (6 H, H Ph), 5.12 (1 H, H1β), 4.54 (1 H, H1α), 3.76-4.15 (8 H), 3.59-3.61 (1 H,), 3.42 (1 H, dt, *J* 8.9 and 0.7, H2), 3.21-3.26 (1 H), 1.05 (9 H, Me).

^{13}C-NMR (101 MHz; $CDCl_3$): δ 154.16 (C quat. Ph), 154.07 (C quat. Ph), 135.61, 134.83, 129.91, 129.85, 127.81 (C Ph), 94.27 (C1β), 94.00 (C1α), 75.17 (C), 74.05 (C), 71.85 (C), 71.10 (C), 71.06 (C), 70.78 (C), 69.11 (C), 65.11 (C), 64.92 (C), 26.93 (C Me), 18.11 (Cq tBu).

MS m/z (EI+) found 441.20 [M+Na] and 859.42 [2M+Na]

1,2,3,4-tetra-O-benzyl-6-O-t-Butyldiphenylsilyl-D-mannopyranose (85)

4.98 mmol of **84** were dissolved in 4,2 mL of dry DMSO; under vigorous stirring, N_2 atmosphere and at 0°C, 12 eq of KOH in powder were added and immediately 25 eq of BnBr were dropped slowly by funnel over 30-40 min, under strong stirring. The reaction was stirred overnight, allowing it to reach room temperature into the same ice bath. The vigorous stirring is a key variable to reach good yields. Then H_2O was added and the aqueous phase was extracted 3 times with DCM. The organic layers were collected and combined, dried over $MgSO_4$ and evaporated; purification by flash chromatography 4x15 cm, from %100 EP, to EP:EtOAc 8:2, gives **85** as a colorless oil (80%).

^{1}H NMR (400 MHz; $CDCl_3$): δ 7.78-7.88 (5 H, H Ph), 7.21-7.56 (25 H, H Ph), 4.96-5.11 (4 H, H Bn and H1β), 4.55-4.75 (4 H, H Bn), 4.52 (1 H, H1α), 4.07-4.14 (3 H, H4, H6, H6'), 4.01 (1 H, d, *J* 2.9, H2), 3.58 (1 H, dd, *J* 9.4 and 3.0, H3), 3.43 (1 H, ddd, *J* 9.5, 3.8 and 2.9, H5), 1.14 (9 H, s, Me)

^{13}C NMR (101 MHz; $CDCl_3$): δ 154.52, 154.18, 154.07, 154.02, 153,44, 153,13 (Cq Ph) 139.1, 138.52, 138.34, 137.6, 136.0, 135.7, 134.0, 129.6, 128.44, 128.39, 128.13, 128.09, 128.04, 127.75, 127.72, 127.65, 127.60, 127.2 (C Ph), 100.0 (1C, C1), 96.8 (1C, C1'), 82.4 (1C, C3), 76.98 (1C, C5), 75.3 (1C, C Bn), 74.8 (1C, C4), 74.3 (1C, C2), 73.9 (1C, Bn), 71.5 (1C, Bn), 70.3 (1C, Bn), 63.4 (1C, C6), 26.8 (1C, Me), 19.4(1C, Cq tBu)

HRMS m/z (EI+): found 801.3587 [M+Na], $C_{50}H_{54}NaO_6Si$ requires 801.3582

1,2,3,4-tetra-O-benzyl-D-mannopyranose (83)

4.03 mmol of **85** were solubilized in 40 ml of dry THF; under N_2 atmosphere, at 0°C and under stirring, 3 eq of TBAF (1M solution in THF) were added, then the reaction was stirred at room temperature, for 4 hours. The solvent was evaporated and the product was purified by flesh cromathography, from EP:EtOAc 8:2 to 7:3, giving **83** as a white sirup (91%).

^{1}H NMR (400 MHz; $CDCl_3$): δ 7.25-7.47 (20 H, Ph), 4.88-5.01 (4 H, Bn), 4.52-4.72 (4 H, Bn), 4.50 (1 H, s H1), 3.91-3.99 (3 H, H2, H4, H6), 3.77-3.83 (1 H, H6'), 3.53 (1 H, dd, *J* 9.4 and 2.8, H3), 3.34 (1 H, ddd, *J* 9.2, 5.6 and 3.2, H5), 2.12-2.19 (1 H, bs OH).

^{13}C NMR (101 MHz; $CDCl_3$): δ 153.90, 153.79, 153.5, 153.0 (1C, Cq Ph), 128.43, 128.37, 128.1, 127.85, 127.79, 127.5 (C Ph), 100.5 (1C, C1), 82.4 (1C, C3), 75.9 (1C, C Bn), 75.3 (1C, C5), 74.9 (1C, C4), 74.10 (1C, C Bn), 74.06 (1C, C2), 71.6 (1C, C Bn), 71.2 (1C, C Bn), 62.6 (1C, C6).

HRMS m/z (EI+): found 563.2413 [M+Na], $C_{34}H_{36}NaO_6$ requires 563.2404

Triflate (86)

Compound **86** was synthesized and used immediately for the next reaction because of its. Glasses, stirring bars and cannula are dried o.n. at 150°C and cooled in vacuum/N_2 to assure the most dry environment for the reactions. Fresh dry solvents and new septa are used to reach the best yields.

Typical procedure involves dissolution of 3,35 mmol of freshly synthesized and clean **83** and 1.7 eq of DTBMP in 16 mL of dry DCM, under N_2 atmosphere in a 100 mL rbf. At -78°C, under stirring, 1,5 eq of a fresh prepared 0.55 M Tf_2O solution in dry DCM, was added by syringe.

After 10-15 min no more starting material is detectable in TLC. Work-up with $NaHCO_{3(aq)}$/DCM or H_2O/DCM had been tested, but unsuccesfully; a fast silica filtration (10cm x 3cm), loading directly the reaction mixture and immediately eluting with 7:3 EP:EtOAc gave the best result.

The solvent was evaporated by rotavapor in a 20°C water bath. The obtained yellow sirup was dried under vacuum in a ice bath; two main factors need to be considered: the solvent (mainly EtOAc) has to be evaporated completely because otherwise it

will interfere with the next reaction, but the degradation of the triflate is directly proportional to the time and the temperature.

Diethyldifluoromethylphosphonate (87)

A solution of 3 eq of HMPA in 16 mL dry THF and a solution of 3 eq $CHF_2PO(OEt)_2$ in 9 mL dry THF were cooled at -78°C under stirring.

When freshly prepared compound **86** was drying under vacuum, 3eq of LDA (2M in THF) were added to the -78°C HMPA solution and then the cold $CHF_2PO(OEt)_2$ solution was added by cannula slowly to the LDA solution. 20 minutes after this addition, the pre-cooled (-78°C) solution of **86** in 16 mL of dry THF was dropped slowly by cannula into the $LiCF_2PO(OEt)_2$ solution.

The reaction was stirred at -78°C for 10 minutes, then it was quenched with 30 mL of saturated $NH_4Cl_{(aq)}$ and the -78°C bath was removed. The mixture was extracted in DCM 4 times, the organic layers were collected, dried over $MgSO_4$ and evaporated.

The obtained oil was loaded on a equilibrated 100g SNAP column and the product purified according to the following procedure:

1cv AcOEt 20% - 30%

6cv 30% - 30%

3cv 30% - 40%

87 was obtained as colorless oil was obtained with 60% yield over two steps.

For these reactions it is important to be fast because of the instability of the triflate, to add the reagents by cannula drop by drop, to use fresh reagents and dry solvents, to pay attention to the temperature when it needs to be constantly at -78°C and do the reaction in a dry environment. Don't leave $LiCF_2PO(OEt)_2$ more than 20-25 minutes before the addition of the triflate.

^{1}-H NMR (400 MHz; $CDCl_3$): δ 7.26-7.37 (20 H, H Ph), 5.00 (2 H, d, *J* 11.0, H Bn), 4.87 (1 H, d, *J* 1.6, H1), 4.77 (1 H, d, *J* 11.7, H Bn), 4.71 (2 H, d, *J* 2.4, H Bn), 4.61 (2 H, HBn), 4.44 (2 H, d, *J* 11.3, H Bn), 4.22-4.28 (5 H, m, 4 CH_2 Et, H5), 3.98 (1 H, dd, *J* 9.2 and 3.1, H3), 3.83 (1 H, dd, *J* 2.9 and 1.9, H2), 3.69-3.74 (1 H, m H4), 2.59-2.74 (1 H, m H6), 2.28-2.44 (1 H, m H6'), 1.33-1.38 (6 H, CH_3 Et).

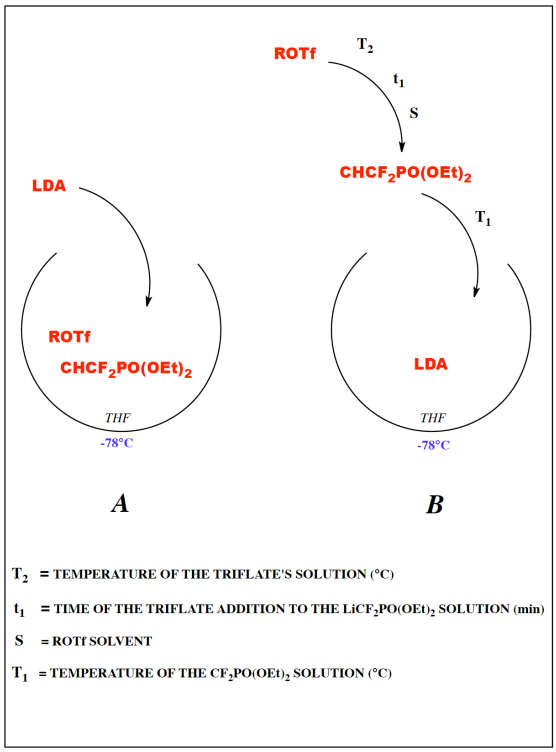
^{13}C-NMR (101 MHz; $CDCl_3$): δ 137.47 (C quat Ph), 128.36, 128.21, 127.93, 127.68 (C Ph), 99.54 (C1'), 96.56 (C1), 80.41 (C3), 77.40 (C4), 74.52 (C2), 72.79, 72.12, 69.57, 68.66 (C Bn), 64.46, 64.39 (CH_2, Et), 51.06 (s, 1C), 35.61 (C6), 16.43 (CH_3, Et).

^{19}F-NMR (377 MHz; $CDCl_3$): δ -110.50 (dddd, *J* = 297.5, 108.0, 34.1, 8.2 Hz, 1F), -113.21 (dddd, *J* = 297.5, 108.8, 29.9, 9.8 Hz, 1F).

^{31}P-NMR (162 MHz; $CDCl_3$): δ 7.22 (app t, *J* = 108.1 Hz, 1P).

HRMS m/z (EI+): found 733.2718 [M+Na], $C_{39}H_{45}F_2NaO_8P$ requires 733.2712

As I previously explained in order to find the right conditions many experiments for this two steps reaction were performed. The two main methods are summarized, in the following figure.



The following table summarizes the most important experiments and it has been showed also in the previous chapter.

4 [mmol]	Tf_2O [eq]	DTBMP [eq]	S/[ml]	Work-up	A/B	LDA [eq]	DIPA/HMPA/nBuLi [eq]	$CHF_2PO(OEt)_2$ [eq]	T_1 [°C]	T_2 [°C]	t_1 [min]	[%]	*
0,234	1,5	1,55	DCM/3	$NaHCO_3$/ice	A	3,5	/	3,6	/	/	/	/	
0,200	1,05	1,05	Et_2O/3,6	Filtration	B	3,5	/	3,5	rt	rt	10	/	
0,183	1,2	1,2	THF/2	/	A	3	/	3	/	/	/	/	*a*
0,081	1,1	1,25	DCM/0,5	H_2O	A	2	/	2	/	/	/	/	
0,185	1,1	1,1	Et_2O/3	Filtration	B	4	/	4	rt	rt	10	6	
0,107	1,1	1,1	Et_2O/3	Filtration	B	2	2 HMPA	2	rt	rt	11	4	
0,181	1,1	1,1	Et_2O/3	Filtration	B	/	4	4	rt	rt	13	7	
0,093	1,2	1,6	DCM/1,1	Silica	A	/	2	2	/	/	/	10	
0,148	1,2	1,6	DCM/1,2	Silica	B	/	3,5	3,5	rt	-78°C	5 – 10	25	
0,148	1,2	2,6	DCM/1,2	Silica	B	/	4	4	rt	-78°C	2	10	*b*
0,148	1,3	1,6	DCM/1,2	Silica	A	/	5	5	/	/	/	8	*c*
0,119	1,2	1,6	DCM/1	Silica	B	/	3,5	3,5	rt	-78°C	20	12	
0,122	1,5	1,6	DCM/1,2	Silica	B	/	3,5	3,5	rt	-78°C	10	14	
0,378	1,5	1,7	DCM/3	Silica	B	/	3,5 + 1,5	3,5 + 1,5	rt	-78°C	30	15	*d*
0,371	1,5	1,7	DCM/3	Silica	B	3,5	3,5 HMPA	3,5	-78°C	-78°C	2	25	*e*
0,404	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	40	
0,487	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	35	13	
0,426	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	33	12	
0,404	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	20	30	
0,404	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	30	
0,835	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	34	
1,332	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	50	
1,852	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	62	
3,354	1,5	1,7	DCM/3	Silica	B	3	3 HMPA	3	-78°C	-78°C	25	60	

A. *$LiCF_2PO(OEt)_2$ in situ production.* **B.** *ROTf added to $LiCF_2PO(OEt)_2$ solution after a time t_1.*

% yield is calculated over two steps

a. I used THF to avoid any work-up, but cooing the solution at -78°C gave a kind of "gelification"-precipitation.no stirring possible

b. Mistake with the cannula addition of the triflate solution: it was not done drop by drop, but fast.

c. Slower elution starting from 9:1 to 7:3 increased the lost of triflate compound in column, and this compromised the yield. 1h reaction didn't change.

d. LDA and difluoro-phosphonate are added in two steps.

e. From now i'm using a new fresh bottle of LDA from Sigma.

Tetraol (88)

1.239 mmol of **87** were dissolved in 13 ml of MeOH. 0.4 eq of $Pd(OH)_2$ (10% on charcoal) were added and while it was stirring, the solution was degassed under vacuum. Then H_2 was added and the reaction was stirred overnight. The reaction was filtered over celite, washed trice with fresh methanol and concentrated under vacuum. The obtained oil was washed with EtOAc, affording **88** with 80% yield.

^{1}H-NMR (400 MHz; CD_3OD): δ 5.02 (1H, d, $J_{1,2}$ = 1.7 Hz, H1), 4.30 (4 H, app quintet, J = 7.2 Hz, CH_2 Et), 4.16 (1H dd, J = 9.5, 9.4 Hz, H5), 3.79 (1H, dd, $J_{2,3}$ = 3.4, $J_{2,1}$ = 1.7 Hz, H2), 3.76 (1H, dd, J = 9.2, $J_{3,2}$ = 3.4 Hz, H3), 3.41 (1H, t, J = 9.5 Hz, H4), 2.77-2.61 (1H, m, H6), 2.32-2.15 (1H, m, H6'), 1.39 (6H, t, J = 7.1 Hz, CH_3 Et).

^{13}C NMR (101 MHz; CD_3OD): δ 94.15(C1), 71.23(C2), 70.4(C4), 70.1(C3), 65.79(C5), 65.74(C5), 64.52,(CH_2 -OEt) 64.46(CH_2 -OEt), 35.5(C6), 15.13(CH_3 -OEt), 15.07(CH_3 -OEt)

^{19}F-NMR (377 MHz; CD_3OD): δ -110.38 (dddd, *J* = 301.0, 110.2, 29.6, 11.6 Hz, 1F), -112.51 (dddd, *J* = 300.9, 110.4, 26.9, 13.3 Hz, 1F).

^{31}P-NMR (162 MHz; CD_3OD): δ 7.20 (app t, *J* = 110.2 Hz, 1P).

HRMS m/z (EI+): found 373.0828 [M+Na], $C_{11}H_{21}F_2NaO_8P$ requires 373.0834

Peracetylated diethyldifluoromethylphosphonate (89)

0.866 mmol of **88** were dissolved in 30 ml of pyridine; 15 ml of Ac_2O and 0,1eq of DMAP were added under stirring and the reaction was stirred overnight. Then the solution was diluted with toluen and concentrated under vacuum. The product was purified by flash cromatographt, from 7:3 to 5:5 EP:EtOAc, affording **89** as a yellow oil 60%.

^{1}H-NMR (400 MHz; $CDCl_3$): δ 6.03 (1 H, d, *J* 1.8, H1), 5.35 (1 H, dd, *J* 10.0 and 3.5, H3), 5.26 (1 H, dd, *J* 3.4 and 2.0, H2), 5.17 (1 H, t, *J* 10.0, H4), 4.35-4.40 (1 H, m, H5), 4.21-4.30 (4 H, m, CH2 Et), 2.18 (3 H, s, Me), 2.17 (3 H, s, Me), 2.07 (3 H, s, Me), 2.01 (3 H, s, Me), 1.35-1.39 (6 H, CH_3 Et).

^{13}C-NMR (101 MHz; $CDCl_3$): δ 170.02, 169.87, 169.73, 168.24 (CO Ac), 90.02 (C1), 68.84 (C3), 68.41 (C2), 64.79 (C4), 66.25 (C5), 64.59 (CH_2 Et), 36.01 (C6), 20.78 (Me), 20.75 (Me), 20.69 (Me), 20.62 (Me), 16.35 (CH_3 Et).

^{31}P-NMR (162 MHz; $CDCl_3$): δ 6.07 (app t, *J* = 104.6 Hz, 1P).

^{19}F-NMR (377 MHz; $CDCl_3$): δ -109.95 (dddd, *J* = 300.8, 105.0, 22.7, 16.7 Hz, 1F), -110.95 (dddd, *J* = 300.8, 104.8, 21.5, 16.1 Hz, 1F).

HRMS m/z (EI+): found 541.1258 [M+Na], $C_{19}H_{29}F_2NaO_{12}P$ requires 541.1257

Peracetylated difluoromethylphosphonate (90)

0,82 mmol of **89** were dissolved in dry toluen and the solvent was evaporated under vacuum; this procedure was repeated three times, adding N_2 instead of air. Then **89** was dissolved in 40 ml of dry DCM. At rt, N_2 atmosphere, 2.3 eq of fresh TMSI were added and the mixture was stirred for 25 minutes. Then the reaction was quenched with 15 mL of MeOH and stirred for 5 min; then 0.5 mL of a 30% solution of $NH_4OH_{(aq)}$ were added to neutralize the HI. The solvent was evaporated under vacuum affording **90** as light orange solid.

^{1}H-NMR (400 MHz; CD_3OD): δ 5.93 (1 H, d, *J* 1.6, H1), 5.34 (1 H, dd, *J* 9.8 and 3.3, H3), 5.29 (1 H, dd, *J* 2.9 and 1.6, H2), 5.12 (1 H, t, *J* 9.9, H4), 4.46 (1 H, t, *J* 9.0, H5), 2.24-2.36 (2 H, H6), 2.13 (6 H, 2Me), 2.05 (3 H, Me), 1.95 (3 H, Me).

^{13}C-NMR (101 MHz; CD_3OD): δ 90.90 (1C, C1), 70.03 (1C, C3), 68.99 (1C, C2), 68.93 (1C, C4), 67.20, and 67.16 (1C, C5), 35.36 and 35.22 (1C, C6), 20.82 (1C, C Me), 20.73 (1C, C Me), 20.65 (1C, C Me).

^{31}P-NMR (162 MHz; CD_3OD): δ 4.19 (t, J = 94.0 Hz, 1P).

^{19}F-NMR (377 MHz; CD_3OD): δ -110.03 (dddd, J = 289.8, 94.2, 27.5, 11.0 Hz), -113.59 (dddd, J = 289.8, 94.3, 26.2, 14.4 Hz).

HRMS m/z (EI+): found 507.0456 [M-H+2Na], $C_{15}H_{20}F_2Na_2O_{12}P$ requires 507.0450

Deprotected difluoromethylphosphonate (91)

0,82 mmol of **90** were dissolved in dry MeOH; under stirring, in N_2 atmosphere 3 eq of MeONa (5M in MeOH) were added. After 10 min in the orange solution was visible a white precipitate. The mixture was loaded in two 15 mL falcons and centrifuged at 3500 rpm for 5 min. The surnatant was collected and the precipitated was washed with 5 mL of MeOH/Et_2O 1:1. After centrifugation the washed solid was dried under vacuum affording **91** as a white solid (40%).

The surnatant obtained from the previous steps was concentrated under vacuum, and re-crystalized in MeOH; the procedure was repeated twice, obtaining 30% of clean and a 20% of bit less clean **91**.

^{1}H-NMR (400 MHz; D_2O): δ 5.04 (1H, d, J = 1.6 Hz, H1β), 4.83 (1H, d, J = 0.9 Hz, H1β), 4.15(1H, t, J = 9.17 Hz, H5), 3.87 (1H, dd, J = 3.4, 0.8 Hz, H2'), 3.85 (1H, d, J = 3.4, 1.8 Hz, H2), 3.76 (1H, dd, J = 9.6, 3.4 Hz, H3), 3.67 (1H, t, J = 8.7 Hz, H5'), 3.56 (1H, dd, J = 9.6, 3.3 Hz, H3'), 3.44 (1H, t, J = 9.7 Hz, H4), 3.37 (1H, t, J = 9.7 Hz, H4'), 2.60-2.43 (1H, m, H6), 2.22-2.05 (1H, m, H6).

^{13}C NMR (101 MHz; D_2O): δ 94.4(C1β), 93.9(C1α), 73.3(C3'), 71.5(C2), 70.9(C3, C5'), 70.6(C4'), 70.3(C4), 70.1(C2'), 67.2(C5), 35.71(C6), 35.57(C6)

^{31}P-NMR (162 MHz; D_2O) 5.53 (1 P, td, *J* 84.8, 10.5)

^{19}F-NMR (377 MHz; D_2O) -110.16 (1 F, dddd, J 278.8, 84.9, 35.6 and 7.0), -112.63 (1 F, dddd, J 278.8, 85.1, 31.8 and 8.7).

HRMS m/z (EI-): found 293.0245 [M-H], $C_7H_{12}F_2O_8P$ requires 293.0243

H_2SO_4-SiO_2 preparation.

5 g of silica were suspended in 25 ml of Et_2O and 1,5 ml of concentrated H_2SO_4 was added. The mixture was stirred for 5 min, then the solvent was evaporated under reduced pressure and the resulting silica was dried for 3 hours at 110°C.

α-Propargyl difluoromethylphosphonate (82)

0.302 mmol of **91** were suspended in 0.5 mL of propargyl alcohol; the mixture was stirred at 60°C for 15 min, then 55 mg of H_2SO_4-SiO_2 were added. After 35 min MS analysis showed only the starting compound; so I added 1 ml of propargyl alcohol and 75 mg of silica. The solution became clean and MS showed only the presence of the final compound.

The reaction mixture was loaded on a 3x5 cm silica column and eluted with EtOAc to wash it from the propargyl alcohol, then EtOAc:EtOH 8:2 and finally MeOH.

The fractions were checked by MS, and those containing **82** have been collected, affording 56 mg of a white solid, containing also silica. According to NMR analysis the compound was not clean.

So I did a second purification using an ion exchange resin; the compound was dissolved in the minimal amount of water and loaded on the Strata SAX column (from Phenomenex).

Elution gradient:

2mL milliQ water

2mL 50mM NH_4HCO_3

2mL 100mM NH_4HCO_3

2mL 150mM NH_4HCO_3

2mL 200mM NH_4HCO_3

8mL 250mM NH_4HCO_3

3mL 350mM NH_4HCO_3

3mL 400mM NH_4HCO_3

3mL 500mM NH_4HCO_3

All the fractions were checked by MS analysis, and **82** was eluted with 50 mM NH_4HCO_3. The fractions were collected and lyophilized affording in 20 mg of a white solid, not enought clean in NMR.

The purification procedure was not good, because I lost most of the compound, while the reaction was giving a very good MS analysis, with the disappearing of the staring compound, affording only the product mass peak.

The solution could be a very short silica filtration to separate the compound from the acidic silica and the propargyl alcohol, followed by HPLC purification.

HRMS m/z (EI-): found 331.0392 [M-H], $C_{10}H_{14}F_2O_8P$ requires 331.0400

α-propargyl-D-mannopyranose (92)

2,78 mmol of D-Mannose were suspended in 20 eq of propargyl alcohol, H_2SO_4-silica (prepared as previously reported) was added (60 mg/mmol) and the reaction was stirred at 65°C. Mannose became soluble after the addition of silica and the mixture was stirred for 3 hours. Then the reaction was filtered on 3x10 cm silica and eluted with AcOET:EtOH (9:1), obtaining **92** (74%).

This reaction was repeated other three times in 1g scale (5,56 mmol) and using fresh prepared silica it was possible reach 70% - 75% yields.

^{1}H-NMR (400 MHz; CD_3OD): δ 4.98 (1 H, H1), 4.28 (2 H, d, *J* 1.3, C$\underline{H_2}$CCH), 3.80-3.86 (2 H, m, H2, H6), 3.73 (1 H, dd, *J* 11.9 and 5.5, H6'), 3.68 (1 H, t, *J* 3.4, H3), 3.63 (2 H, t, *J* 8.1, H4), 3.52 (1 H, app t, *J* 6.5, H5), 2.87-2.91 (1 H, $CH_2C\underline{C}H$), ^{13}C-NMR (101 MHz; CD_3OD): δ 53.99 ($\underline{C}H_2CCH$), 61.73 (C6), 67.41 (C4), 70.97 (C2), 71.49 (C3), 73.97 (C5), 75.25 ($CH_2\underline{C}\underline{C}H$), 79.11 ($CH_2\underline{C}CH$), 98.84 (C1).
HRMS m/z (EI+): found 241.0683 [M+Na], $C_9H_{14}NaO_6$ requires 241.0683

Phosphate (94)

3,3 mmol of **92** were solubilized in 3,4 ml of MeCN, 5 eq Py and 2,2 eq H_2O. At 0°C under stirring 4,4 eq $POCl_3$ were added slowly and the reaction was stirred at 0°C. After 1 hour 38 g of ice were added to quench the reaction. When the ice was melts the pH was neutralized adding 2,5M NaOH solution. The solvents was evaporated by rotavapor and the product was purified by flash cromatography, 7:3 IPA: NH_4OH. Affording compound **94** in quantitative yields. Mass spectra of the product shows that it's not the desired, but a phosphate bridge between position 6 and probably position 4 happened..

^{1}H-NMR (400 MHz; D_2O): δ 5.24 (1 H, H1), 4.54 (2 H, ddd, *J* 3.5, 2.3 and 1.4, CH_2 prop), 4.41-4.50 (1 H, m), 4.39 (1 H, q, *J* 8.6), 4.26 (1 H, d, *J* 10.5), 4.21 (1 H, dd, *J* 3.4 and 1.6), 4.15-4.18 (2 H, m), 3.23-3.24 (1 H, CH prop).

^{13}C-NMR (101 MHz; D_2O): δ 100.36, 79.89, 77.49, 76.45, 70.60, 70.56, 69.10, 69.00, 67.35, 65.38, 65.33, 56.07.

^{31}P-NMR (162 MHz; D_2O): δ -1.87 (s, 1P).

HRMS m/z (EI+): found 303.0248 [M+Na], $C_9H_{14}NaO_6$ requires 303.0240

Difluorenylmethanolphosphate (97)

Difluorenylmethanolchlorophosphate **96** was synthesised solubilising 4,87 mmol (1,5 eq) of $POCl_3$ in 7 mL of dry Py. Under stirring 9,73 mmol (3 eq) of fluorenylmethanol (FmOH) were dissolved in 7 ml dry Py and added by syringe. The solution was vigorously stirred in ice bath for 1 hour and then 3,243 mmol of **92** were solubilized into 7 ml of dry Py and added by syringe at the 0°C solution of **96**, under stirring. The reaction was stirred one hour and then quenched adding H_2O.

The solvent was evaporated and the product was purified by flash chromatography, 9.5:0.5 DCM:EtOH, affording **97,** 45% yields.

^{1}H-NMR (400 MHz; $CDCl_3$): δ 7.73-7.21 (16H, m, H Ph), 5.00 (1H, d, J = 1.3 Hz, H1), 4.34-4.28 (2H, m, H6, CH_2-Fm), 4.22-4.08 (6H, m, H6', CH_2'-Fm, 2CHFm, C$\underline{H_2}$CCH), 3.95-3.92 (1H, m, H2), 3.91-3.85 (1H, m, H4), 3.82 (1H, dd, J = 9.4, 3.2 Hz, H3), 3.63 (1H app d, J = 9.2 Hz, H5), 2.32 (1H, t, J = 2.4 Hz, CH_2CC$\underline{H}$).

^{13}C-NMR (101 MHz; $CDCl_3$): δ 127.88, 127.85, 127.12, 125.14, 125.09, 119.97 (C Ph), 98.62 (1C, C1), 82.97 (1C, $CH_2\underline{C}CH$), 75.20 (1C, $CH_2C\underline{C}H$), 71.45 (1C, C5), 71.21 (1C, C3), 70.33 (1C, C2), 69.80 (1C, d, CH_2 Fm), 69.55 (1C, d, CH_2 Fm), 66.90 (1C, d, C6), 66.59 (1C, C4), 54.46 (1C, $\underline{C}H_2CCH$), 47.78 (1C, CH Fm).

^{31}P-NMR (162 MHz; $CDCl_3$): δ -0.85 (s, 1P).

HRMS m/z (EI+): found 677.1935 [M+Na], $C_{37}H_{35}NaO_9P$ requires 677.1911

α-propargyl-6-phosphate-D-mannopyranose (93)

0,13 mmol of **97** were dissolved in 7 ml DCM and piperidine (0.2 ml/ml DCM) was added under stirring at rt. After 5 min the solvent was evaporated and the solid was extracted in H_2O/DCM 3 times. The first organic fraction was extracted again with water and then the aqueous phase was washed 3 times with DCM. The aqueous phases were collected and evaporated by rotavapor, affording **93**. The NMR showed the presence of piperidine, even leaving the compound overnight under vacuum or after lyophilization. I dissolved 0,13 mmol of **93** in 0,26 ml of water. 2 eq of NaOH 1M solution were added, the solution was stirred and then lyophilized. ^{1}H-NMR showed that piperidine was decreased by a factor of 10. I did not treat further the product with NaOH, because once it will be bound to the protein, by anion exchange it will loose piperidine that will be eliminated with the protein purification.

^{1}H-NMR (400 MHz; DMSO-d6): δ 4.75 (1 H, H1), 4.16 (2 H, qd, *J* 15.6 and 2.3, C$\underline{H_2}$CCH), 3.91-3.96 (1 H, H6'), 3.78 (1 H, dd, *J* 10.5 and 7.9, H6), 3.65 (1 H, d, *J* 9.6, H4), 3.61 (1 H, H2), 3.46 (1 H, t, *J* 2.4, CH_2CC$\underline{H}$), 3.43 (1 H, d, *J* 3.3, H3), 3.29-3.32 (1 H, H5).

^{13}C-NMR (101 MHz; DMSO-d6): δ 99.50 (1C, C1), 80.62 (1C, $CH_2\underline{C}CH$), 74.30 (1C, C5), 71.34 (1C, $CH_2C\underline{C}H$), 71.08 (2C, C3,C2), 66.73 (1C, C4), 64.13 (1C, C5), 64.13 (1C, $\underline{C}H_2CCH$)

^{31}P-NMR (162 MHz; DMSO-d6): δ 2.80 (s, 1P).

HRMS m/z (EI+): found 321.0337 [M+Na], $C_9H_{15}NaO_9P$ requires 321.0346

_N_-tert-Butoxycarbonyl-(_S_)-2-amino-4-amino-butanoic acid (99)

To a solution of N-tert-Butoxycarbonyl-(S)-glutamine (6.00 g, 24.4 mmol) in THF (75 mL) and H_2O (15 mL), PIDA (9.42 g, 29.24 mmol) was added at 0ºC. The solution is stirred o.n. leaving it in the same ice bath. The day after mass spectra (ESI^-) indicated the formation of a product (100%, 217 [M-1]) and complete consumption of starting material. The mixture was evaporated, the residue was dissolved in H_2O (50 mL) and washed with EtOAc (3×40 mL). The aqueous layer was evaporated and lyophilized to give **99** (4.4 g, 83%) as pale yellow solid.

^{1}H NMR (400 MHz, D_2O) δ: 3.96 (m, 1H, CH), 3.04 (at, J = 7.7 Hz, 2H, NH_2CH_2), 2.17-2.04 (m, 1H, CH_2), 1.98-1.89 (m, 1H, CH_2), 1.41 (s, 9H, $C(CH_3)_3$).

^{13}C NMR (400 MHz, D_2O) δ: 178.0 (COOH), 157.9 (CONHBoc), 81.7(($C(CH_3)_3$), 53.6 (Cα), 37.1 (Cβ), 30.0 (Cγ), 29.9 ($C(CH_3)_3$).

ESI^- MS (TOF ES^-)(m/z) calc for [M-H] $C_9H_{18}N_2O_4$= 217.1; found: 217.1.

_N_-tert-Butoxycarbonyl-(_S_)-2-amino-4-azido-butanoic acid (100)

To a solution of sodium azide (13.12 g, 201.16 mmol) in H_2O (60 mL) and CH_2Cl_2 (60 mL) cooled to 0 ºC, triflic anhydride (7.20 mL, 76.8 mmol) was added slowly over 5 min. After 2.5 h of stirring, the phases were separated. Aqueous layer was extracted with CH_2Cl_2 (2×16 mL), and the combined organic layers were washed with saturated aqueous sodium carbonate (20 mL). This triflic azide solution was added to a solution of *N*-tert-butoxycarbonyl-(*S*)-2,4-diaminobutanoic acid **99** (4.40 g, 20.16 mmol), potassium carbonate (4.51 g, 32.70 mmol) and copper sulphate pentahydrate (54.4 mg, 0.22 mmol) in MeOH (130 mL) and H_2O (40 mL) and stirred at rt over the week-end. After 48 h, TLC (1:1:4 H_2O:AcOH:*n*-BuOH) indicated the formation of a product (R_f 0.87), and complete consumption of starting material (R_f 0.5). The organic solvents were evaporated and the remaining aqueous phase was washed with CH_2Cl_2 (3×100 mL), diluted with more H_2O (50 mL) and acidified to pH 2 (1M aqueous HCl). The aqueous phase was extracted with CH_2Cl_2 (3×50 mL), and the combined organic layers were washed with brine (2×80 mL),

dried ($MgSO_4$), filtered and concentrated in vacuum to afford **100** (3.2 g, 66%) as a pale oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 5.21 (br d, *J* = 8.1 Hz, 1H, NH), 4.47-4.29 (m, 1H, CH), 3.46 (at, *J* = 6.7 Hz, 2H, N_3CH_2), 2.26-2.07 (m, 1H, CH_2), 2.06-1.88 (m, 1H, CH_2), 1.46 (s, 9H, $C(CH_3)_3$).

^{13}C NMR (400 MHz, D_2O) δ: 176.1 (COOH), 156.9 (CONHBoc), 80.6 ($C(CH_3)_3$), 60.6 (Cα), 47.8(Cβ), 31.8 (Cγ), 28.2 ($C(CH_3)_3$).

ESI$^-$ MS (TOF ES$^-$)(m/z) calc for [M-H] $C_9H_{16}N_4O_4$= 243.1; found: 243.1.

L-Homoazidoalanine (101)

To a solution of *N*-tert-Butoxycarbonyl-(*S*)-2-amino-4-azido-butanoic acid **100** (3.2 g, 13.09 mmol) in dry CH_2Cl_2 (56 mL), TFA (28 mL) was added and the reaction mixture stirred at rt. After 3.5 h, TLC (1:3 petrol/EtOAc containing 1% TFA) indicated the formation of a major product (R_f 0.07) and complete consumption of starting material (R_f 0.6). Toluen was added and the reaction mixture was concentrated in vacuo to afford (*S*)-2-amino-4-azido-butanoic acid trifluoroacetic acid salt, as a thick yellow oil.

This oil was dissolved in 3mL H_2O, cooled at 0°C and basified with 2.5M NaOH until pH was ~10. Dowex-50WX8 was prewashed with MeOH, acetone, 2M HCl and then milliQ water until the pH was neutral.

This solution was poured onto a column of Dowex-50WX8 (100g conditioned resin). The resin was washed with 500mL of water to remove the TFA salt and then the product was eluted with 5% $NH_4OH_{(aq)}$. The fractions containing homoazidoalanine (ammonium carboxylate) were combined and evaporated under reduced pressure, to give 17 as the free amino acid. This white solid was recrystalized by suspending in ~20mL of EtOH and then adding dropwise H_2O until the solution became clear. When it was clear an additional 10 mL of EtOH was added. The solution is left o.n. in the cold room. The resulting crystals were isolated by filtration and dried under vacuum affording **101** (75%).

^{1}H NMR (400 MHz, D_2O) δ: 3.69 (1H, dd, J=5.6, 7.2, H*α)*, 3.44 (app dt, *J*=6.6, 2.3, 2H, H_γ).

^{13}C NMR (100 MHz, D_2O) δ: 174.2 (*C*OOH), 53.2 (*Cα*), 29.9(*C*β), 47.9 (*C*γ).

HRMS m/z (EI+): found 167.0539 [M+Na], $C_4H_8N_4O_2Na$ requires 167.0545

1) Reaction of Np276 Aha 61-141 (20Å ruler) with α-propargyl-D-mannose-6-phosphate 93

200 µl of diazido-Np276 ($1{,}15*10^{-2}$µmol) 1,2 mg/mL solution were used in this reaction.

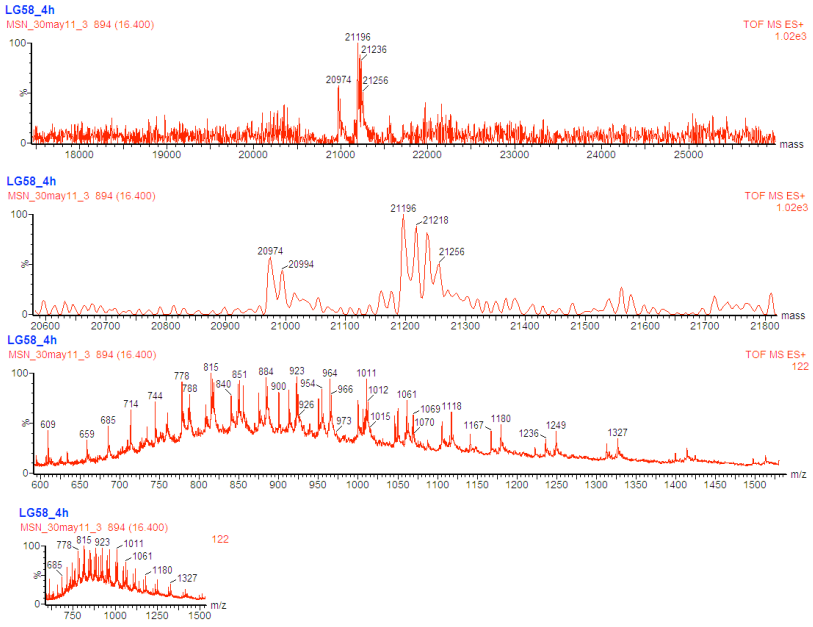
Cu-Complex formation: 163 µl of CuBr 10mg/mL solution were added to 63 µL of a tris-triazolyl amine ligand solution (120mg/mL).

To a solution (1mg/10µL) of 246 eq (123 eq/azide) of **93** in 30 µL of NaPi buffer (50mM, pH 8) were added 200 µL of protein solution (NaPi buffer) and the solution was mixed by vortex.

60 µL of the fresh Cu-ligand solution (126eq of Cu and 170eq of L) were added to this solution, then the solution was mixed by vortex and then mixed gently for 4 hours.

After the reaction time a PD10 column was equilibrated with 8 ml of NaPi buffer. The reaction was diluted to 0,5 mL (+240µL NaPi buffer) and loaded to the column. 1 mL of buffer was added to the column and collected in a vial.

As previously reported mass analysis showed efficient coupling reaction, with a small part of mono-functionalized protein, but total hydrolysis of the phosphate.



21196 = Np276 + 2(11)

20974 = Np276 + (11)

1) Reaction of Np276 Aha 61-141 (20Å ruler) with a-propargyl-mannopyranose-6-phosphate (93)

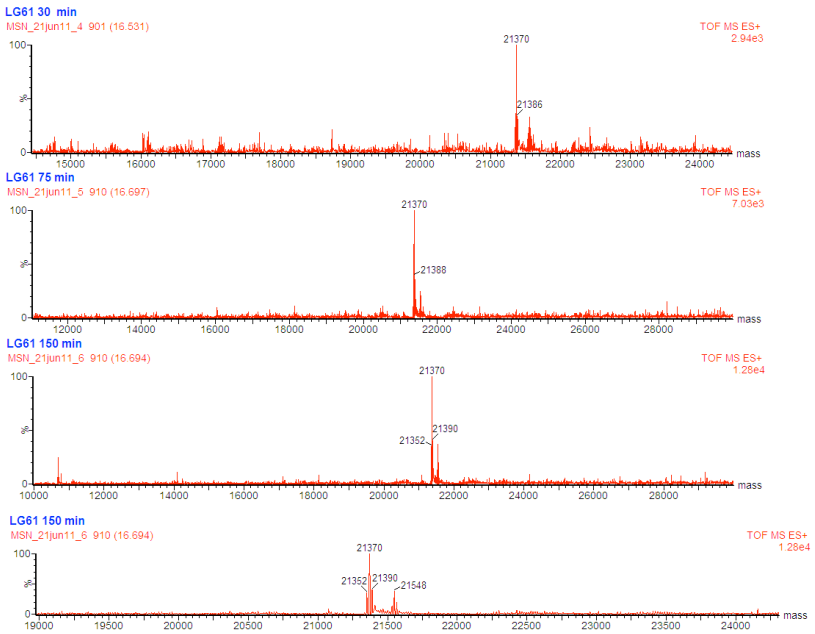
300 µl of diazido-Np276 ($1{,}733*10^{-2}$ µmol) 1,2 mg/ml solution were used in this reaction.

Cu-Complex formation: 188 eq (94 eq/N_3) of CuBr 10 mg/ml solution were added to 133 eq (133 eq/N_3) of a tris-triazolyl amine ligand solution (120 mg/ml).

To a solution (1mg/10µl) of 2000 eq (1000 eq/azide) of 10 (same batch of the last experiment) in 60 µl H_2O were added 300 µl of protein solution (NaPi buffer) and the solution was mixed by vortex.

Fresh Cu-ligand solution (126 eq of Cu and 170 eq of L) was added. The solution was mixed by vortex and then mixed gently. Every 30, 75 and 150 minutes a sample of the reaction was taken and frozen.

Then the samples were purified using size exclusion eppendorf, washing 3 times the volume with fresh buffer. The obtained concentrate solution was analysed, showing a total conversion already after 30 min, without phosphate hydrolysis detectable.



211370 = Np276 + 2(14)

8.4. NMR & biomaterials

^{29}SI-NMR experiments *were performed by Prof. J. Jiménez-Barbero, Centro de Investigaciones Biológicas, Madrid.*

Material fabrication *was performed at Prof. J. R. Jones group, Department of Materials, Imperial College London, England, by dott. L. Russo.*

GPTMS/diamino-PEG hybryds

Hybrids Synthesis. All reagents were supplied by Sigma Aldrich, UK unless otherwise specified. Figure 2 shows a flow chart of the synthesis. Initially 0.500 g of dPEG (M_w 1500 Da) was dissolved in 2 mL of deionised H_2O. This solution was functionalised by addition of an appropriate amount of GPTMS (the molar ratio of dPEG/GPTMS) of 1:2 (dPEG: 0.500 g; GPTMS 0.144 mL) and 1:4 (dPEG: 0.500 g; GPTMS: 0.289 mL). The functionalised dPEG solution was left to mix (2 h) before being added as a sol-gel precursor. A sol was prepared by mixing 592.2 mL of deionised H_2O with 0.197 mL of HCl (1 N) followed by 1.903 mL of tetraethyl orthosilicate (TEOS). The molar ratio of water/TEOS (R ratio) was 4 and HCl 2 N was added to catalyse TEOS hydrolysis at a volume ratio (water/HCl) of 3. The solution was stirred (1 h) to allow hydrolysis of TEOS, before adding the functionalised dPEG solution, which was mixed for a further 1 h. Once the functionalised dPEG solution has mixed for 1 h in the sol, aliquots of the sol (4 mL) were then transferred into screw top teflon moulds (3 cm diameter) for gelling and aging. The gels were aged (3 days at 40 °C) and then dried (8 days at 60 °C).

Hybrid Characterisation. Fourier transform infrared spectroscopy (Perkin Elmer Spectrum 100) was carried out on all samples in the range 4000–250 cm^{-1}, at a resolution of 4 cm^{-1}. The pulse delay was sufficient to prevent saturation. Chemical shifts were referenced externally to TMS at 0 ppm.

Field emission gun scanning electron microscopy (LEO Gemini 1525 FEGSEM) of hybrid nanostructure was carried out on gold coated hybrids to image their nanoscale morphology, with an accelerating voltage of 5 kV and a working distance ~7 mm.

Compression Testing. Compression testing was carried using an Instron 5866 on rectangular samples with height:width ratio of 2:1 (n = 5). A 500 N load cell was used for testing, with a compression extension speed of 0.5 mm min^{-1}. At five

samples (n = 5) of each composite system were measured. The mean and standard deviation were reported. The modulus (taken at the maximum modulus before the onset of yield), yield stress (defined as the end of the linear deformation region) and the corresponding strain at the yield point were also recorded. The yield strength was determined from the cross point of the two tangents on the stress–strain curve around the yield point.

Liquid NMR experiments. Benzyl alcohol was used as internal standard for quantification of the GPTMS-dPEG reaction. This molecule has a high boiling point, it does not have functional groups that could influence the possible reactions, neither at the epoxide moiety nor at the silica network formation, and its ^{1}H-NMR signals do not overlap to those of the reagents. Using the integral of the aromatic protons of benzyl alcohol (H-Ph, m, 5H, d 7.48–7.31 ppm), and referring that to the CH_2 group closest to the silicon atom (H-1, m, 2H, d 0.87–0.60 ppm) and those assigned to the $NH_2CH_2CH_2CH_2$ moiety (H-b, m, 4H, d 2.01–1.89 ppm) of dPEG, the different components were normalised, and the amount of GPTMS and dPEG present at any time in solution was thus estimated.

Hybrid Synthesis. 450 mg of dPEG (0.3 mmol) was dissolved under stirring in 2.55 mL of D_2O. The pH of the solution was adjusted to 6 by adding a DCl 2 N solution in D_2O, then 16 µl of benzyl alcohol (0.154 mmol) were added for the experiments requiring internal standard. 130 µl (0.6 mmol) or 260 µl (1.2 mmol) of GPTMS (respectively for PEG:GPTMS 1:2 and 1:4 experiments) were added and the solution was stirred for 2 h at room temperature, in a sealed vial. After this stage, the vial was transferred to an oven at 40 °C, for 72 h, without stirring. After 2 h at room temperature and after 24, 48 and 72 h at 40 °C, 400 µl of the reaction were collected, diluted adding 100 µl of D_2O, and the resulting solution was analysed by NMR spectroscopy, ^{13}C, ^{1}H and measurements were carried out using a Varian 400 MHz Mercury instrument, operating at a proton frequency equal to 400 MHz; chemical shifts were referenced to methanol (d 3.40 ppm). The 5 mm diameter NMR tube was sealed and maintained closed during the full duration of the experiment. The number of scans was between 160 and 240, depending on the signal-to-noise ratio, and the recycling delay was set to 5 s. ^{13}C-NMR spectra were recorded by using a 200 ppm

spectral width, using between 200 to 1268 scans depending on the signal-to-noise ratio.

NMR studies of GPTMS reactivity as a function of pH

Solutions at pH 2, 5, 7, 9 and 11 were prepared adjusting the pH of D_2O by the addition of DCl (2 M in D_2O) for the acidic solutions, or NaOD (2 M in D_2O) for the basic solutions. 132 µL (0.6 mmol) of GPTMS are added to 0.7 mL of a freshly prepared solution at the title pH (0.85 M), directly in the NMR tubes. The solution was vigorously shaken and analysed by NMR spectroscopy for different reaction times. ^{13}C, ^{1}H, hetero- and homonuclear bidimensional experiments COSY and HSQC measurements were carried out using a Varian 400 MHz Mercury instrument, operating at a proton frequency equal to 400 MHz; chemical shifts were referenced towards methanol (3.40 ppm). The 5 mm diameter NMR tube was sealed and maintained closed during the full duration of the experiment. The number of scans varied between 160 and 240 depending on the signal-to-noise ratio, the recycling delay was 5 s. ^{13}C-NMR spectra were recorded by using a 200 ppm spectral width, using between 200 to 1268 scans depending on the signal-to-noise ratio. The NMR tubes are kept, at room temperature, under gentle mixing and analysed at different reaction's intervals.

D_2O solutions at the desired pH are prepared as described above. In a 1.5 mL Eppendorf®, 132 µL (0.6 mmol) of GPTMS are added to 0,7 mL of a freshly prepared D_2O solution (0.85 M). The Eppendorf® was vigorously shaken and an analytical sample is diluted to 1 mL adding a water/acetonitrile 1:1 solution and then analysed by mass-spectrometry. The Eppendorf® was kept well sealed and the solution analysed at different reaction's time. Mass spectra were recorded on a System Applied Biosystems MDS SCIEX instrument (Q TRAP, LC/MS/MS, turbo ion spray). ESI full MS were recorded on a Thermo LCQ instrument by direct inlet.

9. Abbreviation list

A5P = Arabinose-5-phosphate
Ac_2O = Acetic anhydride
ACN = Acetonitrile
AIBN = Azobisisobutyronitrile
API = Arabinose-5-phosphate isomerase
ATP = Adenosine triphosphate
CD-M6PR = Cation dependent mannose-6-phosphate receptor
CI-M6PR = Cation independent mannose-6-phosphate receptor
CMP = Citydyn monophosphate
CSA = Camphorsulphonic acid
DC = Dendritic cells
DCM = Diclhoro methane
DMAP = 4-Dimethylaminopyridine
DMF = Dimethylformamide
DMSO = Dimethyl sulfoxide
DNA = Deoxyribonucleic acid
DTBMP = 2,6-Di-tert-butyl-4-methylpyridine
ELISA = Enzyme-linked immunosorbent assay
ERT = Enzyme replacement therapy
EtOH = Ethanol
Fm = 9-Fluorenylmethanol
FPLC = Fast protein liquid chromatography
GABA = 4-Aminobutanoic acid
GPTMS = 3-Glycidoxypropyltrimethoxysilane
Haa = Homoazidoalanine
HAS = Human serum albumin
HMBC = Heteronuclear Multiple Bond Correlation
HMPA = Hexamethylphosphoramide
HSC = Hepatic stellate cell
IGF-II = Insulin-like growth factor II
Kdo = 2-Keto-3-deoxy-D-manno-octulosonic acid
LDA = Lithium diisopropylamide
LiHMDS = Lithium bis(trimethylsilyl)amide
LPS = Lipopolysaccaride
LSDs = Lysosomal storage diseases
M6P = mannose-6-phosphate
MeOH = Methanol
NMR = Nuclear magnetic resonance
P-Akt = Phosphorylated Akt
PCR = Polymerase chain reaction
PEG = Polyethylene glycol
PH = Pleckstrin homology
PI3k = Phosphatidylinositide 3-kinase

PIDA = Phenyliodo(III)diacetate
PIPLC = PI-specific phospholipase C
PKB = Protein kinase B
PtdIns = Phosphatidylinositol
PTEN = Phosphatase and tensin homolog
Py = Pyridine
Ru5P = Ribulose-5-phosphate
SEM = Scanning electron microscopy
STD-NMR = Saturation transfer difference nuclear magnetic resonance
TACA = Tumor associated antigen
TBAF = Tetrabutylammonium fluoride
TBAHS = Tetrabutylammonium hydrogen sulfate
TBDPSCl = Tert-butyl diphenylchlorosilane
TEMPO = 2,2,6,6-tetramethyl-1-piperidinyloxyl/hypochlorite
TEOS = Tetraethyl orthosilicate
TFA = 2,2,2-Trifluoroacetic acid
THF = Tetrahydrofuran
TLR = Toll-like receptor
TMSBr = Trimethylsylil bromide
TMSI = Trimethylsylil iodine
iRNA = RNA interference
IC_{50} = Half maximal inhibitory concentration
TEA = Triethylamine
IPA = 2-propanol
PE = Petroleum ether
EtOAc = Ethyl acetate

10. References

1 Gabius H-J. The sugar code: fundamentals of glycoscience. Wiley-Blackwell, 2009
2 Varki A. et al Cell (2006) 126, 841
3 Bishop JR. et al Glycobiology (2007) 17, 23R
4 Schaffer C, et al Proteomics (2001) 1, 248
5 Ritchie GE, et al Chem. Rev (2002) 102, 305
6 Dove A. Nature biotechnol (2001) 19, 913
7 Ohtsubo K, et al Cell (2006) 126, 855
8 Lowe JB, et al Annu. Rev. Biochem. (2003) 72, 673
9 Freeze HH. Nat Rev Genet (2006) 7, 537
10 Dwek RA, Nature Rev. Drug Discov. (2002) 1, 65
11 Van Kooyk Y, et al Nature Immunol. (2008) 9, 593
12 Hakomori, Proc. Natl. Acad. Sci. (2002) 99, 10231
13 Nangia-Makker P, et al Trends Mol. Med. (2002) 8, 187
14 Dube DH, et al Nat. Rev. Drug Discov. (2005) 4, 477
15 Cobb BA, et al Eur. J. Immunol. (2005) 35, 352
16 Brandenburg K, et al Carbohydr. Res. (2003) 338, 2477
17 Ohto U, et al Science (2007) 316, 1631
18 Zimmer SM, Glycobiology (2007) 17, 847
19 Brandenburg K, et al Curr. Top. Med. Chem. (2004) 4, 1127
20 Smith AE, et al Science (2004) 304, 237
21 Raman R, et al Nat Methods (2005) 2, 817
22 Feizi T, et al Curr. Opin. Struct. Biol. (2003) 13, 637
23 Feizi T. Glycoconj. J. (2000) 17, 553
24 Fukui S, et al Nat Biotechnol (2002) 20, 1011
25 Fazio F, et al J Am Chem Soc (2002) 124, 14397
26 Park S, et al Angew Chem Int Ed Engl (2002) 41, 3180
27 Paulson JC, et al Nat. Chem. Biol. (2006) 2, 238
28 Harvey DJ. Mass Spec Rev (2009) 28, 273
29 Kogelberg H, et al Curr Opin Struct Biol (2003) 13, 646
30 Spiro RG. Glycobiology (2002) 12, 43R
31 Wong CH. Carbohydrate-based Drug Discovery. Wiley-VCH Wheineim, 2003

32 Cipolla L. et al Expert Opin. Drug Discov. (2010) 5, (8)
33 Landsteiner, K. Science (1931) 73, 405
34 Morgan WTJ. Some immunological aspects of the products of the human blood group genes. London: Ciba Foundation Symposium, 1959
35 Watkins WM. Biochemistry and genetics of the ABO, Lewis and P blood group systems. Advances in Human Genetics: Plenum Publishing Corp New York, 1980
36 Kabat EA. Clin Pathol (1982) 78, 28l
37 Feizi T, et al Glycoconj J (2000) 17, 439
38 Linhardt RJ. Chem. Ind. (1991) 2, 45
39 Shriver Z, et al Nat Rev Drug Discov (2004) 3, 863
40 Mond JJ, et al Annu Rev Immunol (1995) 13, 655
41 McCool TL, et al Infect Immun (1999) 67, 4862
42 Francis TJr, et al J Exp Med (1930) 52, 573
43 Heidelberger M, J Immunol (1950) 65, 535
44 Nangia-Makker P, et al Trends Mol. Med. (2002) 8, 187
45 Hakomori SI. Cancer Cells (1991) 3, 461
46 Seeberger PH, et al Nat Rev Drug Discov (2005) 4, 751
47 Boltje TJ, et al Nature Chemistry (2009) 1, 611
48 Seeberger, P. H. Chem Soc Rev (2008) 37, 19
49 Davies BG. J.C.S., Perkin (2000) 14, 2137
50 Barkley A, et al Chem Eur J (2001) 7, 555
51 Nicolau KC, et al Angew Chem Int Ed (2001) 40, 1577
52 Hanessian S. Total Synthesis of Natural Products: The “Chiron” Approach. Pergamon: Oxford, 1983
53 Hollingsworth, et al Chem Rev (2000) 100, 4267.
54 Hirschmann RF, et al Acc Chem Res. (2009) 42, 1511
55 Velter I, J Carb Chem (2006) 25, 97
56 Cipolla L, et al Curr. Org. Synthesis (2005) 2, 153
57 Meutermans W, et al Chem Med Chem (2006) 1, 1164
58 Le GT, et al. Drug Discov Today (2003) 8, 701
59 Nicolaou KC, et al Angew Chem Int Ed (2001) 40, 1576
60 Capozzi G. et al Bioorg. Med. Chem. Lett. (2002) 12, 2263
61 Peri F, et al J Chem Soc Perkin Trans I (2002) 5, 638
62 Araújo AC, et al Carbohydr Res (2008) 343, 1840
63 Araújo AC et al Eur J Org Chem (2008) 635
64 Ernst B et al Nature Rev Drug Discov (2009) 8, 661
65 Koeller KM et al Nature Biotechnol (2000) 18, 835

66 Lillelund VH et al Chem Rev (2002) 102, 515
67 Gruner SAW, et al Chem Rev (2002) 102, 491
68 Gravier-Pelletier C et al Curr Org Synth (2007) 4, 1
69 Arjona O et al Chem Rev (2007) 107, 1919
70 Compain P, Martin O. Iminosugars: from synthesis to therapeutic applications. John Wiley & Sons Ltd. The Atrium Southern Gate: England 2007
71 Simanek EE et al Chem Rev (1998) 98, 833
72 Lynn M et al J Infect Dis (2003) 187, 631
73 Piazza M et al J Med Chem. (2009) 52, 1209
74 Christ, W.J.; Rossignol, D.P.; Kobayashi, S.; Kawata, T.; *U.S. Patent.*, 1999, 5, 935.
75 Von Itzstein et al M Nature (1993) 363, 418
76 Lew W et al Curr Med Chem (2000) 7, 663
77 Cipolla L et al Current Med. Chem. (2011) 18, 830. Gabrielli L. et al Current Drug Targets, (2012) 13, 1458
78 Holst O FEMS Microbiol. Lett. (2007) 271, 3
79 Raetz C.R.Het al Annu. Rev. Biochem. (2002) 71, 635
80 Holst O Trends Glycosci. Glycotech. (2002) 14, 87. Holst, O.; Brade, H. In: Chemical structure of the core region of lipopolysaccharides. Bacterial Endotoxic Lipopolysaccharides, Morrison DC & Ryan JL, eds, CRC Press, Boca Raton, FL, **1992**; Vol.*1*, pp. 135–170.
81 Nikaido H Microbiol. Mol. Biol. Rev. (2003) 67, 593
82 Tokuda H et al Biochim. Biophys. Acta (2004) 1693, 5
83 Kamio Y et al Biochemistry (1976) 15, 2561
84 Ulevitch R.J. et al Annu. Rev. Immunol. (1995) 13, 437
85 Kim H.M.et al Cell, (2007) 130, 906. Zimmer S.M. et al Glycobiology, (2007) 17, 847. Ohto U et al Science (2007) 316, 1632
86 Angus D.C et al Crit. Care Med. (2001) 29, 1303
87 Christ, W.J.; Rossignol, D.P.; Kobayashi, S.; Kawata, T. Substituted lipopolysaccharides useful in the treatment and prevention of endotoxemia.
U.S. Patent 5935938, August 10, 1999. Lynn M et al J. Infect. Dis. (2003) 187, 631
88 Alexander C et al Trends Glycosci. Glycotechnol. (2002) 14, 69
89 Viriyakosol S et al J. Biol. Chem. (2001) 276, 38044. Poltorak A et al Proc. Natl. Acad. Sci. U.S.A. (2000) 97, 2163. Lien E et al J. Clin. Invest. (2000) 105, 497. Muroi M et al Infect. Immun. (2002) 70, 3546
90 Brandenburg K et al Curr. Top. Med. Chem. (2004) 4, 1127. Montminy S.W et al Nat. Immunol. (2006) 7, 1066. Brandenburg. K et al Carbohydr. Res. (2003) 338, 2477. Mata-Haro V et al Science (2007) 316, 1628

91 Kusumoto S et al J. Endotoxin Res. (2003) 9, 361
92 Means T.K et al Cytokine Growth Factor Rev. (2000) 11, 219
93 Cipolla L Nat. Prod. Rep. (2010) 27, 1618
94 Angata T et al Chem. Rev. (2002) 102, 439. Schauer R Zoology (2004) 107, 49
95 Levin D. H. et al J. Biol. Chem. (1959) 234, 2532
96 Ghalambomr M et al J. Biol. Chem. (1966) 241, 3207
97 Reynolds CM et al Biochemistry (2009) 48, 9627
98 Frirdich E et al J. Biol. Chem. (2005) 280, 27604
99 Gronow S et al J. Endotoxin Res. (2001) 7, 3
100 Yethon J.A et al Curr. Drug Targets Infect. Disord. (2001) 1, 91
101 King JD et al Innate Immunity (2009) 15, 261
102 Meredith TC et al J. Bacteriol. (2005) 187, 6936
103 P. H. Ray, J. E. Kelsey, E. C. Bigham, C. D. Benedict and T. A. Miller, ACS Symposium Series 231, ed. L. Anderson and
F. M. Unger, American Chemical Society, Washington, D. C., 1983, pp. 141–169.
104 Dotson GD et al J. Biol. Chem. (1995) 270, 13698
105 Wu et al J. Biol. Chem. (2003) 278, 18117
106 Goldman RC et al J. Bacteriol. (1985) 163, 256
107 Goldman RC et al J. Biol. Chem. (1988) 263, 5217. Mohan S et al J. Bacteriol. (1994) 176, 6944. Rocchetta HL et al Microbiol. Mol. Biol. Rev., 1999, 63, 523.
108 White KA et al J. Biol. Chem. (1997) 272, 16555
109 Belunis KE et al J. Biol. Chem. (1992) 267, 18702
110 Sperandeo P et al Res. Microbiol. (2006) 157, 547
111 Welch RA Proc. Natl. Acad. Sci. U. S. A. (2002) 99, 17020. Tzeng YL et al J. Biol. Chem. (2002) 277, 24103
112 Meredith TC et al J. Biol. Chem. (2003) 278, 32771
113 Baetman A Trends Biochem. Sci. (1997) 22, 12
114 Sommaruga S et al Biochem. Biophys. Res. Commun. (2009) 388, 222
115 Dotson GD et al J. Biol. Chem. (1995) 270, 13698
116 Taylor PL et al J. Biol. Chem. (2008) 283, 2835. Tello M et al ChemBioChem (2008) 9, 1295. Jeffery CJ et al Biochemistry (2001) 40, 1560. Straus D et al Proc. Natl. Acad. Sci. U. S. A. (1985) 82, 2272. Lolis E et al Biochemistry (1990) 29, 6619
117 Rose I Adv. Enzymol. (1975) 43, 491
118 Walsh C Enzymatic Reaction Mechanisms, W. H. Freeman & Co., San Francisco, 1979, pp. 585–599.
119 Airoldi C et al Chem. Eur. J. (2010) 16, 1897

120 Yep A et al Bioorg. Med. Chem. Lett. (2011) 21, 2679
121 Kohen A et al Eur. J. Biochem. (1992) 208, 443
122 Airoldi C Journal of Carbohydrate Chemistry (2010) 29, 30
123 Sekiguchi et al Proc. Natl. Acad. Sci. USA (1967) 58, 2315
124 Dangerfield EM et al. Carb. Res. (2010) 345, 1360
125 Bigham EC et al *J.* Med. Chem. (1984) 27, 717
126 Chunikhin KS et al Russian Chemical Reviews (2010) 79 371
127 Engel R Chemical Reviews (1977) vol 77**,** n 3
128 Unger et al Carb Res (1978) 67, 349
129 Chambers RD J. Chem. Soc. Commun (1990) 1053
130 Hu J Chem. Commun. (2009) 7465
131 Kaboudin B Beilstein Journal of Organic Chemistry (2006) 2, n 4
132 Chlebowski JF The Journal of Biological Chemistry (1974) 249, n 22 7192
133 Bigham EC Journal of Medicinal Chemistry (1984) 27, 717
134 Gary-Bobo M et al Curr. Med. Chem (2007) 14, 2945
135 Vidil C et al Eur. J. Org. Chem. (1999) 447
136 Vidal S et al Eur J Org Chem. (2000) 3433
137 Ambrose MG J. Org. Chem. (1983) 48, 674
138 Bertowitz et al J Org. Chem. (1993) 58, 6174
139 Dangerfield EM et al Carb. Res. (2010) 345, 1360
140 Al-Horani RA Tetrahedron (2010) 66, 2907
141 Zhao M et al J. Org. Chem. (1999) 64, 2564
142 Berkowitz DB et al Org. Lett. (2004) 6, 4921
143 Bhaskar PT et al Dev Cell. (2007) 12, 487
144 Gonzalez E Cell Cycle (2009) 8, 2502
145 Benedetti V Mol Cancer Ther. (2008) 7, 679
146 O'Reilly KE et al Cancer Res. (2006) 66, 1500
147 Miyagi T Glycoconj J (2012) 29, 567
148 Le GT Drug Discovery Today (2003) 8, 701
149 Li Q. Expert Opin. Ther. Patents (2007) 17, 1077
150 Cipolla L et Carb. Res. (2010) 345, 1291
151 Rong SB J. Med. Chem (2001) 44, 898
152 Tropper FD et al Synthesis (1992) 618
153 Kannan T et al Bioorg. Med. Chem. Lett. (2001) 11, 2433
154 Györgydeák Z et al J. Carb. Res. (1995) 286, 85
155 Bertini R et al Proc. Natl. Acad. Sci. U.S.A. (2004) 101, 11791. Allegretti M et al J. Med. Chem.

(2005) 48, 4312
156 Rong SB et al Med. Chem. (2001) 44, 898
157 Park et al Nat. Biotechnol. (2006) 24, 1581
158 Steinman R J. Med. (2001) 68, 160
159 Akira S. Adv. Immunol. (2001) 78, 1. Schnare M et al Nat. Immunol. (2001) 2, 947
160 Beutler B et al Annu. Rev. Immunol. (2006) 24, 353
161 Kannan T Bioorg. Med. Chem. Lett. (2001) 11, 2433
162 Winzler C J. Exp. Med. (1997) 185, 317
163 Granucci F J. Immunol. (2003) 170, 5075
164 Leblais V et al Circ. Res. (2004) 5, 1183. Oudit G J. Mol. Cell. Cardiol. (2004) 37, 449
165 Liu P et al Nature Rev. Drug Discov. (2009) 8, 627
166 Liu D et al Cancer Res. (2009) 69, 7311
167 Kornfeld S et al Annu. Rev. Cell Biol. (1989) 5, 483. Kornfeld S et al Annu. Rev. Biochem. (1992) 61, 307
168 Schmidt B et al J. Biol. Chem. (1995) 270, 14975.
169 Kang JX et al., PNAS U.S.A. (1998) 95, 13671.
170 Kandror KV et al. Biochem J (1998) 331, 829
171 Oka Y et al J Biol Chem (1985) 260, 9435
172 Kornfeld S, Ann. Rev. Biochem. (1992) 61, 307. Munier-Lehmann A et al Biolchem. Soc. Trans. 1996, 24, 133. Dahms NM et al., Biolchem. Soc. Trans. 1996, 24, 136.
173 Ni X et al Istol Histopathol. (2006) 21, 899. Nyakjaer et al J. Biol. Chem. (1998) 141, 815
174 Dahms NM et al., Glycobiol. (2008) 18, 664.
175 Gary-Bobo M et al Curr. Med. Chem. (2007) 14, 2945
176 Reddy ST et al J. Biol. Chem. (2004) 279, 38658
177 Byrd GC et al J. Biol. Chem. (2000) 275, 18656
178 Varki A et al J. Biol. Chem. (1980) 255, 10847. Varki A et al J. Biol. Chem. (1983) 258, 2808. Tong PY et al J. Biol. Chem. (1989) 264, 7962
179 Distler J et al J. Biol. Chem. (1991) 266, 21687
180 York SJ et al J. Biol. Chem. (1999) 274, 1164
181 Mammen M et al Angew Chem Int. (1998) 37, 2755
182 Fei X et al Med. Chem. Lett. (2008) 18, 3085
183 M.K. Christensen et al., J. Chem. Soc. Perkin Trans I 1994, 1299.
184 Tong PY et al J Biol. Chem. (1989) 264, 7970
185 Diestler JJ et al J Biol. Chem. (1991) 266, 21687
186 Bardy RO Annu. Rev. Med (2006) 57, 283
187 Lee K et al Glycobiology (2003) 13, 305

188 Berthe ML et al Eur. J. Cancer. (2003) 39, 635
189 Beljaars L et al Hepatology (1999) 29, 1486
190 Zangh J et al Glycoconjugate J. (2003) 19, 423
191 Ye Z et al Biochem. (2005) 44, 4466
192 Ruiz-Cabello J et al *NMR Biomed.* (2009) 24, 114
193 Vetting et al Protein Science (2007) 16, 755
194 Lu W et al Carb. Res. (2005) 340, 1213
195 Mukhopadhy B et al Hepatology (1999) 29, 1486
196 Kartha KPR et al Tetrahedron (1997) 53, 11753
197 Shaihk N et al Mol. Div. (2010) DOI 10.1007/s11030-010-9281-2. Yeoh K et al Carb. Res (2009) 344, 586
198 Sakakura A et al Org. Let. (1999) 7, 1999
199 Zangh J et al Glycoconjugate J. (2003) 19, 423
200 St. Hilaire PM et al Rapid Commun. Mass Specrom. (1998) 12, 1475
201 Andruszkiewicz R et al Synth. Commun. (2004) 34, 1049
202 Hench LL et al Science (2002) 295, 1014
203 Jones JR et al Biomaterials (2006) 27, 964
204 Novak BM Adv Mater (1993) 5, 422
205 Jones J. R Acta Biomater 2012, http://dx.doi.org/10.1016/j.actbio.2012.08.023.
206 Valliant EM et al Soft Matter (2011) 7, 5083
207 Vallet-Regi Curr Nanosci (2006) 2, 179. Tian D et al J. Polym. Sci. Part A: Polym. Chem. (1997) 35, 2295. Catauro MG et al J. Mater. Sci. (2003) 38, 3097. Chen Q et al J. Sol-Gel Sci. Technol. (2000) 19, 101. Pereira MM et al Mater. Sci. Mater. Med. (2005) 16, 1045. Ren Let al Biomaterials (2002) 23 4765–4773. Rhee S Biomaterials (2004) 25, 1167
208 Mahony O et al Advanced Functional Materials (2010) 20, 3835
209 Judeinstein P et al J. Mater. Chem. (1996) 6, 511. Novak MB Adv. Mater. (1993) 5, 422. Mackenzie J et al J. Sol-Gel Sci. Technol. (1996) 7, 151–161.
210 Liu et al MATERIALS SCIENCE & ENGINEERING C-MATERIALS FOR BIOLOGICAL APPLICATIONS (2012) 32, 707
211 Poologasundarampillai G et al Mater. Chem. (2010) 20, 8952
212 *Dalton Trans.*, 2009, 9146–9152
213 Robertson MA et al Journal of Sol-Gel Science and Technology (2003) 26, 291
214 Innocenzi P et al Chem. Mater. (1999) 11, 1672. Innocenzi P et al Chem. Mater. (2000) 12, 3726
215 Journal of the Ceramic Society of Japan (2011) 119, 387
216 Innocenzi P et al Journal of Sol-Gel Science and Technology (2005) 35, 225
217 Martin RA et al Trans R Soc London A (2012) 1422

218 Davis SR J Non-Cryst Solids (2003) 197
219 Gizdavic-Nikolaidis MR et al J Non-Cryst Solids (2007) 353, 1598
220 Costa SH et al J Mater Sci (2008) 43, 494
221 John Coates in: Encyclopedia of Analytical Chemistry R.A. Meyers Ed, Interpretation of Infrared Spectra, A Practical Approach, John Wiley & Sons Ltd, Chichester, 2000, pp. 10815–10837
222 Schottner G Chem. Mater. (2001) 13, 3422
223 Dave BC et al J Sol-Gel Sci Technol (2004) 32, 143
224 Zheludkevich ML et al Surf Coat Technol (2006) 200, 3084
225 Wu KA et al J Polym Sci Part A: Polym Chem (2006) *44*, 335-342.
226 Que W et al *Appl Phys A* (2001) 73, 171
227 Zhang JL et al J Electron Mater (2008) 37, 135
228 Gill I et al Tibtech (2000) 18, 282
229 G. Dubois, W. Volksen, R.D. Miller, In: *Dielectric Films for Advanced Microelectronics*, (Eds: M.Baklanov, K. Maex, M. Green), Wiley, New- York, **2007**, Chap 2.
230 Wang QG et al Macromol. (2003) 36, 5760
231 Liu J et al J Adhes Sci Technol (2006) 20, 277
232 Kasten LS et al Progr. Org. Coat. (2003) 47, 214–224
233 Tenzplin M et al Spiess Adv. Mater. (1997) 9, 814
234 Davis SR et al J Non-Cryst Solids (2003) 197
235 Mammeri F et al Mater. Chem. (2005) 15, 3787
236 Snodgrass JM et al Acta Mater. (2002) 50, 2395
237 Maly M et al Mol. Simul. (2008) 34, 1215
238 Innocenzi P et al Dalton Trans. (2009) 9146
239 Innocenzi P et al J. Ceram. Soc. Jap. (2011) 119, 387
240 Menaa B et al Chem. Mater. (2007) 19, 1946
241 Innocenzi P et al Chem. Mater. (2001) 13, 3635
242 Hoebbel D et al J Sol-Gel Sci Technol. (2001) 21, 177
243 Alonso B et al Chem. Mater. (2005) 17, 3172
244 Oliver MS et al J Sol-Gel Sci Technol (2010) 55, 360
245 Innocenzi P et al Chem. Mater. (2001) 13, 3635
246 Airoldi C et al (2011) 12, 719-27.
247 Wiegand I et al Nat Protoc. 3, 163.
248 Datsenko KA et al Proc Natl Acad Sci. (2010) 97, 6640
249 Babak Beilstain journal of Org. Chem. (2006) 2, 4
250 Molecular Operating Environment; Chemical Computing Group Inc. http://www.chemcomp.com/.
251 Halgren, T. A. J. Comput. Chem. 1999, 20, 720–729.

252 Baumann, H.; Öhrman, S.; Shinohara, Y.; Ersoy, O.; Choudhury, D.; Axen, A.; Tedebark, U.; Carredano, E. Protein Sci. 2003, 12, 784793.
253 Rong, S. B.; Hu, Y.; Enyedy, I.; Powis, G.; Meuillet, E. J.; Wu, X.; Wang, R.; Wang, S.; Kozikowski, A. P. J. Med. Chem. 2001, 44, 898–908.

11. Acknowledgments

Well, these three years are gone very quickly but I'm happy to say they are a very beautiful part of my life.

And this is sincerely thanks to my Prof., Laura Cipolla, a perfect maestro and a perfect boss for me, that make me fooling in love with science even more that I was before.

Prof. Nicotra, because not many people are able to tech you a lot, just during a coffee break.

Laura Toni Russo is one of the greatest people I ever met and she simply saved my life.

Prof Jesus Jiménez-Barbero, because it is great to work and discuss about everything with him.

Thanks to Prof. B.G. Davis, for the opportunity to work in his group and because that months in Oxford were an intense and beautiful experience that taught me a lot.

My parents and Manuel; I do not have words to thanks them, but I can sincerely say that for me they are the best parents and brother in the word, and without them I could not be here, writing this thesis.

Thanks to my lab-mates, Bini for our romantic nights, Antonella, Alice, Cristina for all the funny moments.

Thanks to Mitul, Justin, Matt, Guillolme, Emilie, Bala, Macarena, Alba, Charles and all the BGD group because it was fantastic work, spent time and have party with you.

Thanks to my friends, that were always with me in these years, even if I was not with them.

If I have done a PhD is also thanks to all the people that I met in my life and increased in me the passion and the love for science: maestro Marino, dott Galliani, prof. Rindone, prof Alexander Kros and Hana.

During my life I often took crazy and foolish choices, but I am a very lucky person, because if I could come back I would exactly do the same choices that led me here.