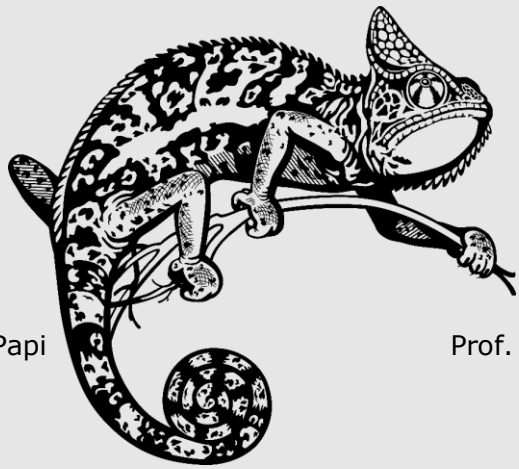
DOTTORATO DI RICERCA IN
SCIENZE CHIMICHE

CICLO XXXIII

COORDINATORE Prof. PIERO BAGLIONI

Structurally constrained mimetics of tumor-associated MUC1 antigens

Settore Scientifico Disciplinare CHIM/06



Dottorando
Dott. Francesco Papi

Tutore
Prof. Cristina Nativi

Anni 2017/2020

DOTTORATO DI RICERCA IN
SCIENZE CHIMICHE

CICLO XXXIII

COORDINATORE Prof. PIERO BAGLIONI

Structurally constrained mimetics of tumor-associated MUC1 antigens

Settore Scientifico Disciplinare CHIM/06

Dottorando
Dott. Francesco Papi

Tutore
Prof. Cristina Nativi

Coordinatore
Prof. Piero Baglioni

Anni 2017/2020

Table of Contents

Abstract

Summary

Part I - Introduction

Chapter 1: Immune system 10

1.1 An overview 10

1.2 Immune system 11

1.3 Adaptive immune response 13

1.3.1 B cells 14

1.3.2 T cells 15

1.3.3 Antigen-presenting cells (APCs) 16

1.3.4 Hapten-carrier effect 17

Chapter 2: Cancer immunotherapy 18

2.1 Cancer therapy: an overview 18

2.2 Active and passive immunotherapy 19

2.3 Tumor antigens (TAs) in cancer immunotherapy 20

Chapter 3: Mucin 1-type O-glycans 22

3.1 Mucin 1 22

3.2 Mucin-related antigens 23

3.2.1 The α-Tn antigen 24

3.2.2 The TF antigen (Core 1) 25

3.2.3 The STn antigen 25

3.3 Mucin-type O-glycans vaccine candidates *26*

Chapter 4: Tn antigen analogues **28**

4.1 Tn antigen analogues: an overview *28*

4.2 O-Glycosidic bond modification *29*

4.2.1 C-glycosidic bond 29

4.2.2 S and Se-glycosides 30

4.3 Tn analogues containing unnatural amino acids *31*

4.4 Constrained Tn analogues *33*

4.5 Thr vs. Ser *34*

Chapter 5: A locked α-TnThr mimetic **36**

5.1 A powerful inverse electron-demand [4+2] cycloaddition *36*

5.2 Multivalent presentation of the TnThr mimetic 13 ..... *39*

5.2.1 Superparamagnetic iron oxide nanoparticles 40

5.2.2 Dextran-based single chain polymer nanoparticles 41

5.2.3 A tetravalent synthetic cyclopeptide carrier - RAFT 42

Chapter 6: Aim of the project **44**

Part II - A TnThr mimetic

Chapter 7: A breast cancer vaccine candidate **46**

7.1 Introduction *46*

7.2.1 An optimized synthesis of the TnThr mimetic 48

7.2.3 Structural characterization of CRM_{197} glycoconjugates 61

7.2.4 Binding of mime[4]CRM to anti-Tn antibody 64

7.2.5 *In vitro:* activation of human dendritic cells (DCs) 65

7.2.6 *In vivo:* evaluation of the therapeutic effect of mime[4]CRM by using TNBC transplanted model 66

7.2.7 Anti TnThr mimetic specific antibody response elicited by mime[4]CRM immunization 73

7.2.8 Isolation of mime[4]CRM-specific T cells. 77

7.3 Conclusions *77*

7.4 Experimental section *79*

7.4.1 General methods 79

7.4.2 Synthesis 89

Chapter 8: Silica nanoparticles decorated with a TnThr mimetic as tool for solid state NMR (SSNMR) investigation 107

8.1 Introduction *107*

8.2 Results and Discussion *109*

8.2.1 Synthesis of TnThr-Silica nanoparticles (TnThr-SiNPs) 109

8.2.2 SSNMR and ICP-MS characterization of GSiNPs 112

8.3 Conclusions *113*

8.4 Experimental section *115*

8.4.1 General methods 115

8.4.2 Synthesis 116

Part III - A TFThr mimetic

Chapter 9: Interaction of a TFThr tumor-associated antigen mimetic with Gal-3 120

9.1 Introduction *120*

9.2 Results and Discussion *121*

9.2.1 Synthesis of a TFThr mimetic 121

9.2.2 NMR and ITC binding studies of 43 vs. Gal-3 CRD 126

9.2.3 X-ray of the Gal-3 CRD-TFThr mimetic complex 131

9.3 Conclusions *134*

9.4 Experimental section *136*

9.4.1 General methods 136

9.4.2 Synthesis 139

Part IV - An STnThr mimetic

Chapter 10: Synthesis of a novel STnThr mimetic 147

10.1 Introduction *147*

10.2 Results and Discussion *149*

10.2.1 Synthesis of an STnThr mimetic 149

10.2.2 Recognition of the STnThr mimetic by *Influenza* neuraminidase 158

10.4 Experimental section *162*

10.4.1 General methods 162

10.4.2 Synthesis 163

10.4.3 H1N1 Neuraminidase recognition 175

Part V - Conclusions

Abbreviations 180

Bibliography 186

Abstract

Mucins are high molecular weight glycoproteins largely expressed by endothelial cells. They represent a physical protection for organs, against harmful species.[1] When normal mucin become tumoral, a pathological oversimplification of the glycosidic portions occurred and saccharides, normally hidden, become exposed.[2] These neo-glycans, called *mucin-associated antigens*, are interesting targets for the development of therapeutic cancer vaccines because of their almost exclusive presence on cancer cells.[3] In particular, the structurally simple antigens associated to mucine-1 (MUC-1), namely antigen Tn, TF and STn, have been proved to be involved in tumour escape and metastasis. Recent studies showed a clear relationship between the presentation of these antigens and the amino acid involved in the glycosidic linkage; in particular the Thr residue induces a better presentation of the antigen with respect to Ser.[4] In this context, the design of analogues and mimetics of these antigens offer an intriguing strategy to overcome the two main drawbacks affecting the natural derivatives: (a) low immunogenicity and, (b) reduced metabolic stability.[5] In this project we aim at synthesizing structural mimetics of the antigenic TnThr, TFThr, STnThr and at employing these new compounds to decorate multivalent systems to study their *in-vitro* and *in-vivo* properties.

Summary

In this thesis the following topics are addressed:

Part I - Introduction: A general introduction to the basic concepts of immunology (**Chapter 1**), cancer immunotherapy (**Chapter 2**) and tumor-associated antigens (**Chapter 3**), followed by the state-of-art on analogues of mucin-related antigens (**Chapter 4 and Chapter 5**) and the aim of the project (**Chapter 6**).

Part II - TnThr mimetic: Following the recent progress, in **Chapter 7** the therapeutic potential of the TnThr mimetic is highlighted. A vaccine candidate against breast cancer was developed by the conjugation of the TnThr mimetic to the immunogenic carrier protein CRM_{197}, *in silico*, *in vitro* and *in vivo* assays were performed.

In **Chapter 8** an experimental methodological application is described: the TnThr mimetic was grafted onto silica nanoparticles and solid-state NMR (SSNMR) was performed, paving the way for new perspective in the characterization of biomolecules-coated nanoparticles.

Part III - TFThr mimetic: In **Chapter 9**, the glycomimetic strategy was applied for the synthesis of a TFThr mimetic; binding studies of the mimetic with Galectin 3 are discussed.

Part IV - STnThr mimetic: In **Chapter 10,** the synthesis of a novel STnThr mimetic is reported. To investigate the effect of the structurally constrained galactoside moiety, the STnThr mimetic was checked as potential substrate for Influenza neuraminidase N1.

Part I

Introduction

CHAPTER

1

Immune system

1.1 An overview

The overall function of the immune system is to prevent or reduce infections and protect against diseases, through a complex cascade of processes called immune response. The immune system apparatus is constituted of cells and molecules which participate to a concerted and finely tuned response against harmful (immunogenic) entities.[6]

In this context, antigens are defined as any substances that possess antigenic sites (epitopes), subjected to the recognition by the immune system; those antigens capable of eliciting an immune response are called immunogens.[7] In mammalians, immune defense can be subclassified according to the different role played by each component.

1.2 Immune system

Classically, the immune system is divided into two main categories (**Figure 1.2.1**), innate and adaptive, although these distinctions are not mutually exclusive:

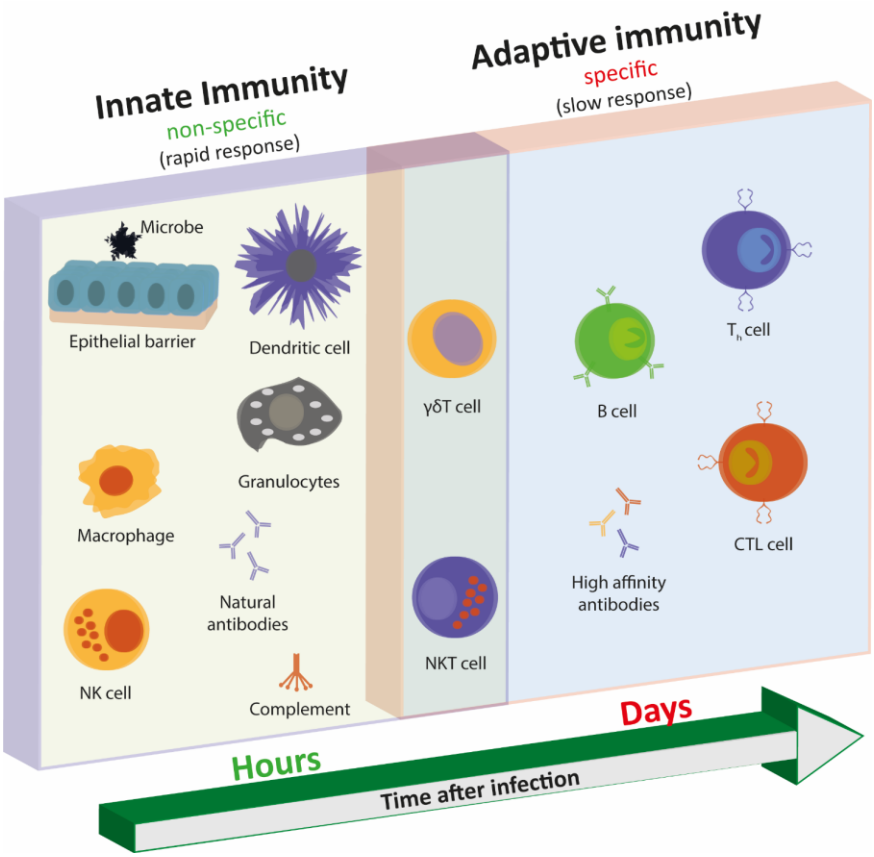


Figure 1.2.1 Innate *vs* adaptive immunity

Innate immunity (also known as *natural*) constitutes the first set of responses, it is present in almost all living organism and includes physical barriers, leukocytes, innate lymphoid cells, soluble proteins, bioactive small molecules.[8] The receptors expressed by the cells of the innate system are called primitive or pattern recognition receptors (PRRs) and

the entities recognized are referred as pathogen-associated molecular patterns (PAMPs). The activation of the innate immunity is fast but lacks specificity.

Adaptive immunity (also referred as *specific* or *acquired*) constitutes the second set of responses and comes into play when innate defense is not enough. The species involved in these type of immunity are: antibodies (Abs), cytokines, antigen-presenting cells (APCs), lymphocytes (B and T cells), major histocompatibility complex (MHC). Adaptive immunity activation is slower but more specific and creates immunological memory.

Innate and adaptive immunity should not be considered as two distinct entities, indeed they synergistically respond to endo and exogenic threats.[9]

Immune system activation and protection is also a central matter in tumor events. Adaptive immunity might represent a powerful "*weapon*" for cancer-bearing patients to fight cancer, unfortunately too often the activation of an innate immune response against tumors is lacking or ineffective. [10–13]

Although a vast effort is devoted to develop immunotherapeutic strategies against cancer, no active cancer immunotherapies are currently available. In this framework, this manuscript is focused on the synthesis and development of nature-inspired determinants to assemble potential therapeutic cancer vaccines.

1.3 Adaptive immune response

The hallmark in the adaptive immune response of vertebrates is the clonal expansion of lymphocytes that consists in the rapid expansion and differentiation of few precursor cells to millions of specialized B and T cells.[14]

Two main types of acquired immunity must be considered (**Figure 1.3.1**): passive and active immunity.

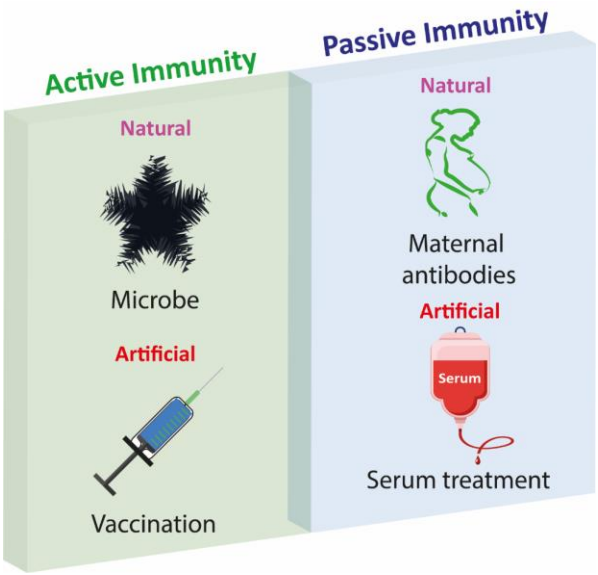


Figure 1.3.1 Passive and active immunity

Passive immunity: the transfer of active humoral immune species (antibody-related) is mediated by biological fluids, as it happens during life-saving serum treatments or in the case of maternal antibodies transmission. This process is also called *adoptive transfer* and typically confers a non-persisting immunity.

Active immunity: developed after the exposure to a microorganism or an immunogenic exogenous substance. Lymphocytes evolve in response to the antigenic "non-self-species" and the clonal expansion takes place. This process can induce immunological memory. Different types of active immunological responses can take place, depending on the nature of the recruited lymphocytes.

1.3.1 B cells

B cells are the main actors in humoral immunity, they induce antibodies formation and act as APCs through the recognition, internalization and processing of antigens.[15]

B cells can recognize antigenic determinants through specific receptors (B Cells Receptors, BCR) and through two different processes:

T cell-dependent activation: antigens activate B cells in cooperation with T cells; no humoral response is triggered in organisms lacking T cells. These antigens, known as thymus-dependent (TD) antigens are processed and presented to T cells with MHC class II molecules. Immunogenic proteins are examples of TD antigens.[16]

T cell-independent activation: antigens stimulate B cell proliferation and antibody production in the absence of MHC class II-restricted T cell help; these antigens are classified as thymus-independent (TI) antigens.[17] Typically highly repetitive molecules, such as polysaccharides of bacterial cell walls, are examples of TI antigens.[18]

B cells activated by TD and TI antigens undergo two differentiation pathways: a) to plasma cells and Abs production for immediate protection

and and/or b) to memory B cells for a long-lasting protection.[18,19] Noteworthy, polysaccharide antigens elicit memory B cells that are phenotypically distinct from those elicited by protein antigens.[18]

1.3.2 T cells

T cells are the main actors of cell-mediated adaptive immunity. Antigens cannot be recognized directly by T cell receptors (TCR) unless processed by APCs and presented with an MHC molecule.[20]

T cells can be divided in four main categories: a) effector T cells, b) memory T cells, c) natural killer T cells (NKT) and d) Tγδ cells.

Widespread and broadly studied, effector and memory T cells represent the conventional adaptive T lymphocytes and they can be further subclassify according to the expression of the surface glycoproteins CD4 or CD8 that confer them the ability of recognizing different classes of MHC molecules: MHC class I for CD8 and MHC class II for CD4.[15,20]

Effector - *helper CD4$^+$ T cells* (T_H) are involved in both innate and adaptive response. T_Hs assist other lymphocytes by promoting: *a)* proliferation and differentiation of B cells into plasma cells and memory cells, *b)* differentiation of cytotoxic CD8$^+$ T cells and c) activation of macrophages through secretion of cytokines.

Effector - *cytotoxic CD8$^+$ T cells* (CTLs) have the ability to lyse cells exposing short peptide antigens-MHC I complexes.[21] CTLs are responsible for immune surveillance against pathogen and cytolysis of somatic cells; they are also involved in the production of cytokines influencing B cells and APCs activity.[22]

Effector - *regulatory T cells* (T_{REG}) act to suppress immune response, thereby maintaining homeostasis and self-tolerance. T_{REG} can express both $CD4^+$ and $CD8^+$.[23]

Memory T cells (T_M) remain long term after an immunological response, they can quickly be converted into effector cells (T_H or CTLs) in case of a subsequent contact with the same antigen. T_Ms ensure a rapid and robust immune response and are the hallmark of vaccination strategies.[24,25]

1.3.3 Antigen-presenting cells (APCs)

Even though, B cells and some species of the innate immune system can work in a "*stand-alone*" way by recognizing antigens that have not been formerly processed, antigen presentation represents a crucial issue to trigger an adaptive immune response (T cells, *vide infra*).

Cells deputed to antigens presentation are called antigen-presenting cells (APCs). Dendritic cells, macrophages, and B cells are APCs for T cells, while follicular dendritic cells mainly are APC for B cells.[26]

APCs are characterized by: *a)* the presence of receptors able to uptake antigens, *b)* the capacity to internalize and process antigens, *c)* to synthesize and express MHC class I and II molecules, *d)* to express adhesion molecules that promote APC-T cell interaction and *e)* to express co-stimulatory molecules (*i.e.* cytokines).[15] Antigen presentation is therefore defined as the expression of processed antigenic molecules on the surface of APCs, in association with MHC class I molecules, when the presentation is to CTLs cells or with MHC class II molecules, when the antigen is presented to a T_H cell.

MHC class I can accommodate short peptides (8-9 aa) originating from cytosolic or nuclear locations while MHC class II can accommodate larger peptides (> 10 aa)[27]. The phenotype of the immune cells depends on the different MHC compartment implicated in the immune response.

1.3.4 Hapten-carrier effect

So far, we dealt with the principal mechanisms responsible for eliciting an immune response, generally induced by high molecular weight immunogens (>10 000 Da) like: proteins, carbohydrates, nucleic acid or lipids. Noteworthy, also small molecules, intrinsically not immunogenic, can elicit an immune response if covalently linked to a large immunogenic carrier. These low-molecular-weight compounds are called “haptens” and the hapten-carrier derivative is called “conjugate”. The conjugate is thus an immunogenic entity, and induce an immune response not only with respect to the carrier but also to the hapten, in a recognition process known as *hapten-carrier effect*.[15,20] More interestingly, the antibodies produced are able to recognize the hapten also as single entity.[28] This strategy, successfully used in bacterial vaccine development such as for *Haemophilus influenzae* type B,[29] is the basis of the development of antiviral and antitumoral carbohydrates vaccines.[30]

CHAPTER

2

Cancer immunotherapy

2.1 Cancer therapy: an overview

As reported by the World Health Organization "Cancer is a large group of diseases that can start in almost any organ or tissue of the body when abnormal cells grow uncontrollably, go beyond their usual boundaries to invade adjoining parts of the body and/or spread to other organs. The latter process is called metastasizing and is a major cause of death from cancer".

Conventional cancer treatments are:

Surgery: the physical removal of cancer, very effective for local tumor but inadequate to large invasive tumoral masses and hematological cancers; after surgical extirpation of a tumor mass metastatization and infection are the main risks.

Chemotherapy: a therapy based on anti-cancer drugs, called chemotherapeutic agents, administered to reduce/eliminate tumoral

mass. Adverse side effects related to the lack of specificity of chemotherapeutic agents are often the causes of chemotherapy's failure.

Radiation therapy: based on ionizing radiation used to control or kill cancer cells by damaging their DNA and to prevent tumor recurrence. Radiotherapy (like chemotherapy) is based on the higher growth rate of cancer- compared to healthy cells; side effects are dose related, in addition, not all kinds of tumors are responsive to this therapy.

A combination of tumor treatments can be prescribed to obtain a synergistic effect and best outcomes, but up to date general guidelines do not exist and a poor prognosis is often registered also with a combination of therapies.

A new *tool* to fight tumor, *cancer immunotherapy*, it has recently become popular.[31] This frontier approach aims at boosting tumor-bearing immune system to fight tumor and improve immunosurveillance.[13]

2.2 Active and passive immunotherapy

Immunotherapy of cancer, although in its infancy, has already drawn a great deal of interest as potential personalized treatment of cancer. Conventional treatments' failure and the need of alternatives to non-specific and invasive therapies,[32] have been the driving forces. The "mission" of immunotherapy is boosting patients' native defenses and stimulating the immune system to detect and selectively tackle tumor cells.

Cancer immunotherapy can be distinguished into two main classes:[33]

Active immunotherapy which relies on setting up an immune response to kill cancer cells by targeting specific tumor antigens (TAs) expressed on their surface, or realizing immunomodulatory agents that enhance the immune responsiveness, without a specific target, such as checkpoint inhibitors.[34–36]

Passive immunotherapy consists of the administration of exogenous substances like Abs or lymphocytes (adoptive cell transfer) designed to target tumor cells.

Given the aim of this manuscript, active immunotherapy involving TAs will be treated in more detail.

2.3 Tumor antigens (TAs) in cancer immunotherapy

In TAs-based immunotherapy, the targets for the immune system are antigens exclusively expressed on cancer cells (i.e. *tumor-specific antigens*, TSAs), or poorly expressed on normal tissues (i.e. *tumor-associated antigens*, TAAs).[37,38] These tumor antigens are altered self-molecules like: proteins, glycoproteins, glycolipids or carbohydrates that differentiate cancer cells from normal cells and can be exploited for preventive, therapeutic and diagnostic purposes.[39,40]

Vaccines based on TAs constitute an interesting area of immunotherapy designed to stimulate a "personalized" antitumor immune response.[41–43] A subclassification of immunotherapy between prophylactic (prevention) and therapeutic (treatment) must also be taken into account for cancer vaccines.[44,45] Cancer vaccines might be used: a) as mono-therapy, b) in combination with conventional treatments in a multi-therapy setting, to

increase its effectiveness and reduce the recurrence of metastatic sites, or as c) preventive tools.[46] Advantages of TAs preventive and therapeutic strategies are a minimal toxicity compared to other treatments (chemo and radio-therapies) combined with an high specificity of action. Examples of approved therapeutic cancer vaccines are TheraCys® (urothelial carcinoma), PROVENGE (metastatic castration-resistant prostate cancer), IMLYGIC® (melanoma).[44]

TAs can be of high interest also from a diagnostic point of view, in that they might be biomarkers in screening for early detection and/or to classify the tumor in order to decide on the most appropriate pharmacological therapy.[40,47] Indeed, whatever the purpose, the choice of the right TAs might make the difference in terms of therapy outcome and it represents a hard and painstaking challenge.

For this reason, the next chapter will be focused on the glycoprotein *mucin-1* (MUC1) overexpressed in many adenocarcinomas and which is characterized by the presence of specific TAAs.

CHAPTER

3

Mucin 1-type O-glycans

3.1 Mucin 1

The unbalance of the glycosylation machinery is one of the main factors involved in carcinogenesis. On cancer cells aberrant glycosylation takes place, including glycan truncation, neo-glycan expression, and increase/decrease in the density of glycan epitopes.[48] In adenocarcinomas, a common aberration occurring is the over-simplification of the glycosylation pattern of mucins. Mucins are a group of high molecular weight glycoproteins, expressed by various epithelial cell types, having a central role in maintaining homeostasis and providing organs' protection.[1] Over the years, among the 21 mucins identified in humans, MUC1 has been the most extensively studied due to its involvement in cancer and chronic inflammatory diseases.[49]

MUC1 is a type I transmembrane glycoprotein with a unique extracellular domain consisting of a *variable number tandem repeat* (VNTR) of 20 amino acids (HGVTSAPDTRPAPGSTAPPA) where *serine* (Ser) and *threonine*

(Thr) are potential sites for O-glycosylation and represent 25% of residues.[50,51] Tumor-associated MUC1, (TA)MUC1, is overexpressed, under-O-glycosylated and exposes specific and structurally simple saccharide epitopes called *mucin-related antigens*, that have been identified in almost all human epithelial adenocarcinomas[52] and in some hematological malignancies.[53,54] Normally hidden, mucin-related antigens are part of a broader family of TAAs called *tumor-associated carbohydrate antigens* (TACAs). TACAs are potential diagnostic molecular markers and attracting targets to develop carbohydrate-based cancer vaccines.[52]

3.2 Mucin-related antigens

The protein backbone of healthy mucins is decorated with complex branched α-O-linked oligosaccharides, this essential post-translational

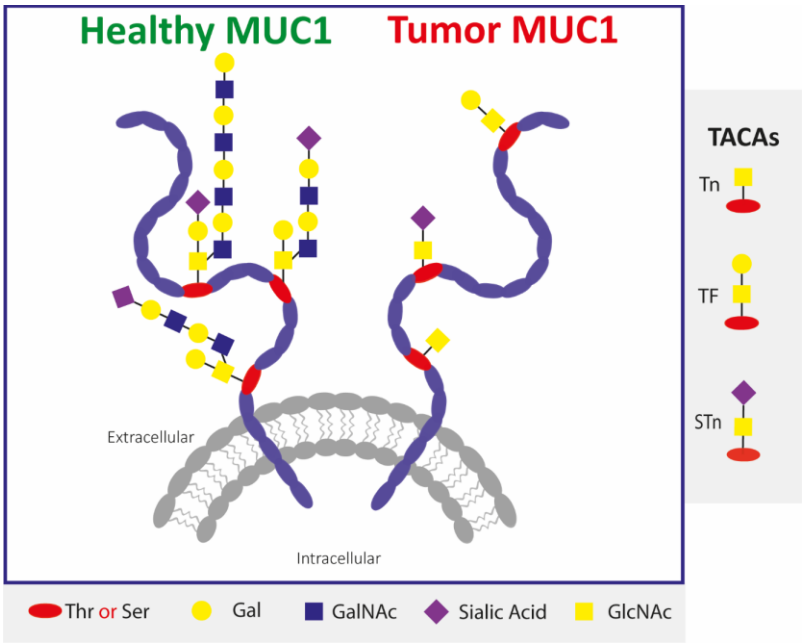
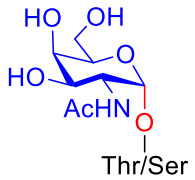


Figure 3.2.1 Differences between healthy MUC1 and tumor MUC1

modification begins with the transfer of an N-acetylgalactosamine (GalNAc) onto the hydroxyl groups of a Ser or Thr residue. Misregulation in glycosyltransferases' activity occurs in tumor cells[55] and leads to the expression of (TA)MUC1, characterized by simple and truncated carbohydrates (**Figure 3.2.1**). The TACAs most extensively studied are: *α-Tn* (or Tn), *TF* (T or Core 1) and *STn,* whose expression has been related to tumor aggressiveness and tendency to metastasize.[39] The elucidation of TACAs chemical structures has been essential to accelerate the research on the development of new "*tools*" to fight cancer.[55,56]

3.2.1 The α-Tn antigen

The *Thomsen-nouveau* (α-Tn) antigen was firstly discovered by Dausset in 1959,[57] in a patient with hemolytic anemia (Tn syndrome). Almost 20 years later, the Tn antigen was associated to malignant tumors.[2,58–60] Structurally, the Tn antigen (**Figure 3.2.1.1**) is a GalNAc moiety α-O-linked to a Ser or Thr residue (GalNAc-α1-O-Ser/Thr) and is the common precursor of the eight different core structures of mucin-type O-glycans. In cancer cells, deficiency in glycosyltransferases, limits Tn glycosylation and saccharides extensions thus making this antigenic residue exposed.

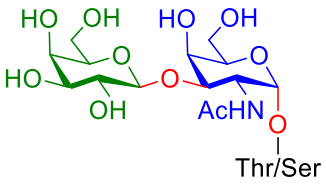


Tn antigen

Figure 3.2.1.1 Structure of *Thomsen-nouveau* antigen

3.2.2 The TF antigen (Core 1)

The *Thomsen–Friedenreich* (TF) disaccharide epitope, also referred as Core-1 or T antigen (**Figure 3.2.2.1**), is a Galβ1→3GalNAc disaccharide derived from the elongation of the Tn antigen with an additional Gal residue mediated by the action of β1,3-galactosyltransferase (Core 1 β3GalT, T-synthase). The TF antigen is usually masked under a layer of other branched and complex carbohydrate residues, the loss of this "cover" makes it visible and antigenic. The expression of TF antigen plays a crucial role in cancer development as a result of its involvement in adhesion processes and tumor cells proliferation through the interaction with specific lectins such as galectin-3 (Gal 3).[61]



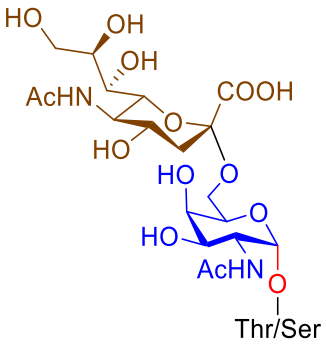
TF antigen

Figure 3.2.2.1 Structure of *Thomsen–Friedenreich* antigen

3.2.3 The STn antigen

The sialyl-Tn (STn) antigen (**Figure 3.2.3.1**) formation is mediate by α2,6-sialyltransferases (ST6GalNAc), which is up-regulated in cancer cells and is responsible for the aberrant sialylation of the Tn antigen. Due to their similar structure and pathways of synthesis, STn and Tn antigens are always over-expressed simultaneously in tumor cells. STn up-regulation is related to tumor growth and metastasis in breast, gastric and prostate

cancers.[62,63] Recognition by lectins (selectins, siglecs, *etc.*), even in this case, has been shown to be crucial in cancerous processes.



STn antigen

Figure 3.2.3.1 Structure of sialyl-Tn antigen

3.3 Mucin-type O-glycans vaccine candidates

In summary, the Tn antigen is highly expressed in 90% of breast carcinomas, [39] some bladder, cervix, ovary, colon, lung, stomach, and prostate tumor and it is poorly or absent on normal adult tissues. TF and STn, overexpressed on cancer cells, have also been proved to be involved in tumor progression.[61–63] These features combined with their simple structures, makes these three determinants ideal targets for the development of anti-cancer vaccines. Over the years, the potential of Tn, TF and STn antigens as therapeutic targets has widely been investigated,[56,64,65] however, two main drawbacks affecting all TACA-based vaccines have slowed down their development: poor immunogenicity and a reduced metabolic stability.

Poor immunogenicity is intrinsic in TACAs' nature; two main strategies are generally followed to boost TACAs immunological response : *a) the* co-administration with adjuvants, *b) the* conjugation to immunogenic carriers (**Chapter 7**).[52,66]

Reduced metabolic stability of TACAs can be ascribed to the glycosidic linkages, unstable *in vivo* and hydrolyzed by endogenous glycosidases. This instability is responsible for a low bioavailability and concurs to their low immunogenicity.

In this framework, synthetic chemistry can play a role of utmost importance as will be described in detail in the next chapter.

In particular, we focused on Tn antigen due to its simple structure and to the great interest of the scientific community on the possible development of Tn-based vaccines.

CHAPTER

4

Tn antigen analogues

4.1 Tn antigen analogues: an overview

Tn antigen, like all *O*-glycosides are degraded *in vivo* by glycosidases. In recent years, many strategies to elude chemical and enzymatic degradation have been explored. Chemical synthesis offers a possible way out by designing hydrolysis-resistant Tn antigen *analogues*.[5]

Before discussing the strategies pursued, let us consider some basic characteristics of a good antigen analogue:

a) it ought to compete with the native TACA for the same biological target; *b)* it must be able to break patient's immune tolerance, eliciting an effective immune response and the development of antibodies directed against the natural counterparts; *c)* it should present a non-native glycosidic linkage and, consequently, an improved stability and higher bioavailability; d) last but not least, a good analogue shall maintain the

pharmacophoric conformation of the native epitope and eventually, should be structurally simpler and synthetically convenient.

Hereafter, some non-native, Tn-inspired antigens and their presentation to the immune system are reported.

4.2 O-Glycosidic bond modification

A straightforward modification of the labile α-O-glycosidic linkage featuring the Tn antigen to obtain hydrolytically stable analogues is represented by the non-classical isosteric replacement of the oxygen with carbon, sulfur, or selenium.

4.2.1 C-glycosidic bond

Beau *et al.* paved the way for the synthesis of C-glycoside analogues of the Tn epitope, and through a SmI_2-mediated C-glycosylation, the TnSer C-analogue **1** (**Figure 4.2.1.1**), as building block for peptide synthesis, was prepared for the first time.[67] Few years later, a Wittig-Horner tin-mediated C-glycosylation variant for the synthesis of **1**, was proposed by Schmidt *et al.*[68]

Following this line, the α-C-GalNAc derivative **2** showing a ketone residue (**Figure 4.2.1.1**), was prepared by Nicotra and co-workers[69] and covalently linked to the *ovalbumin* (OVA) T cell epitope peptide through an oxime residue (**3** - **Figure 4.2.1.1**). Derivative **3** represents an example of a fully synthetic vaccine candidate, stable to acids, bases, and enzymatic hydrolysis.[70]

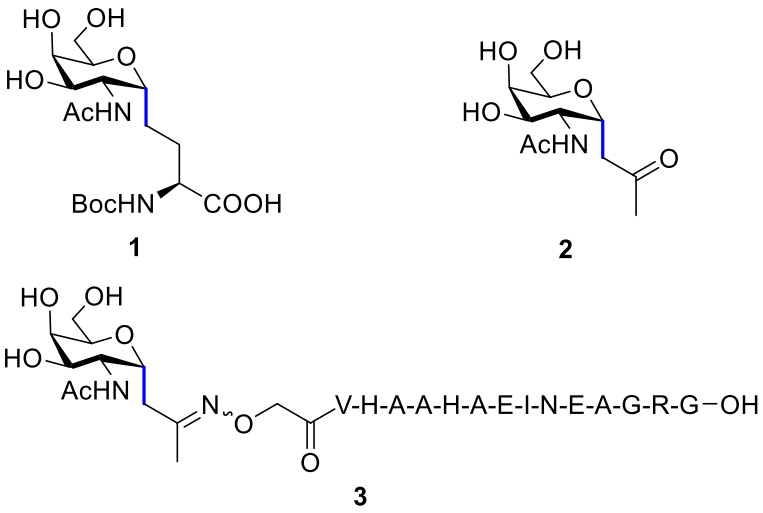


Figure 4.2.1.1 C-glycosidic analogues of Tn antigen

4.2.2 S and Se-glycosides

The first metabolically stable S-analogue of the TnSer determinant has been reported by Spadaro *et al.* (compound **4** - **Figure 4.2.2.1**). Analogue **4** was linked to the immunogenic lipopeptide tripalmitoyl-S-glycerylcysteinylserine (Pam3CysSer), directly or through decorated calix[4]arene scaffold, to develop totally synthetic vaccines.[71,72]

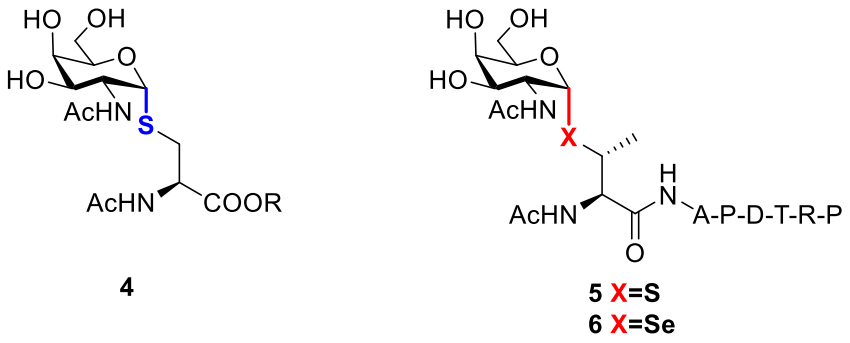


Figure 4.2.2.1 Structures of S (**4**) and Se-glycosidic analogues **5** and **6**

O/S (**5**) or O/Se (**6**) substitutions at the glycosidic linkage (**Figure 4.2.2.1**) connecting the GalNAc residue to the MUC1 repeated, immunodominant, hexapeptide fragment (Ala-Pro-Asp-Thr-Arg-Pro) has been recently reported by Corzana *et al.* The TnThr derivative **5** bearing the S-glycosidic linkage was conjugated to gold nanoparticles and resulted in a significant humoral immune response in mice (BALB/c mice).[73]

4.3 Tn analogues containing unnatural amino acids

A different approach relies on the use of unnatural amino acid that could hamper glycosidases activity thus increasing the hydrolytic stability of the resulting TACA analogue.

The efficient synthesis of the propargylated-α-O-GalNAc derivative **7** (**Figure 4.3.1**), used in a Cu(I) catalyzed conjugation with an azido serine, has been reported by Brimble *et al*.[74] Compound **7** was used as building block in the synthesis of antifreeze glycopeptides (AFGPs) analogues.

Dumy, Renaudet and co-workers reported the conjugation of four clustered GalNAc residues on a cyclic decapeptide scaffold through non-native oxime bond ligation (**8** - **Figure 4.3.1**). The regioselectively addressable functionalized template (RAFT), contains two Pro-Gly residues as β-turn inducers.[75] This scaffold, upon the insertion of a T-epitope peptide, has served for the construction of synthetic vaccines candidates[76,77] able to elicit antibodies in immunized mice (BALB/c), recognized by the Tn epitope expressed on human tumor cells.[75]

Inspired by the X-ray structure of the SM3 monoclonal antibody complexed with a MUC1 minimal epitope, a Tn analogue characterized by

the non-natural α-methylserine scaffold **9** (**Figure 4.3.1**), has been reported by Corzana, Peregrina *et al.*[78] Although the peptide backbone of this vaccine candidate is more resistant to enzymatic degradation than the natural counterpart, its immunogenicity does not exceed that of the threonine analogue.[79]

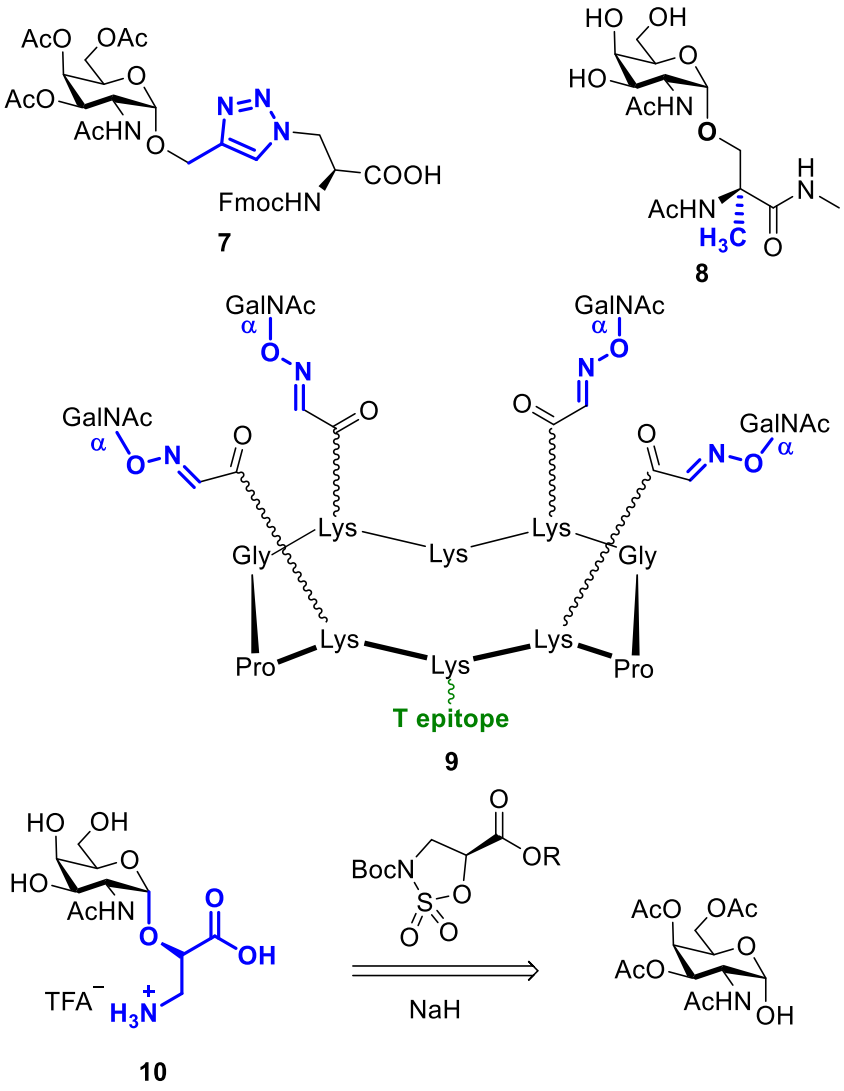


Figure 4.3.1 Structures of Tn analogues containing unnatural amino acids

Interestingly, Peregrina *et al*. developed a strategy to obtain Tn analogues based on the stereoselective ring-opening of hindered and cyclic sulfamidates of (S)-isoserine and (S)-α-methylserine. The α-methylserine derivative **9** and the new analogue α-GalNAc-isoSer **10** (**Figure 4.3.1**) were obtained.[80]

4.4 Constrained Tn analogues

In the quest for enhanced stability and improved immunogenicity, some examples of bicyclic constrained Tn analogues which preserve the $^4C_1$ chair conformation of the native antigen, have been reported. Indeed, a reduced conformational flexibility may potentially reduce the entropic penalty associated with the binding of the antigen to the receptor.

Nativi *et al.* proposed a chemo-, regio-, and stereoselective inverse electron-demand Diels-Alder cycloaddition (**Figure 4.4.1**) between galactal **11** and thio-β-ketoester **12** (see **Chapter 5** for more details) for the synthesis of the structurally constrained TnThr mimetic **13**.[81,82]

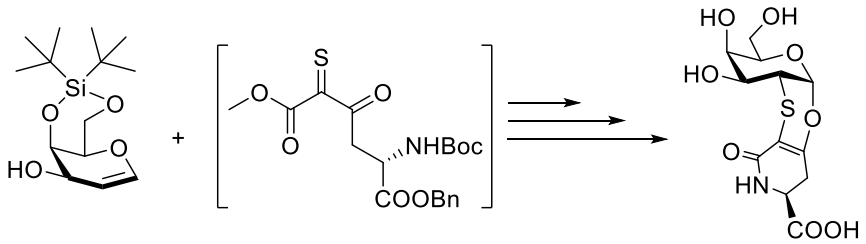
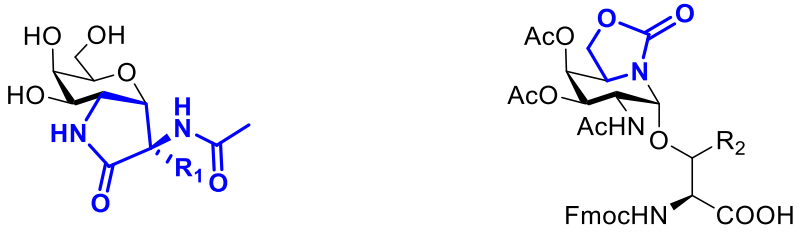


Figure 4.4.1 A structurally constrained TnThr mimetic **13**

Following this approach, in 2013 Peregrina *et al.* reported a totally stereo controlled C-Michael addition for the synthesis of the conformationally restricted C-glycoside analogues of TnSer **14** (**Figure 4.4.2**), in which the

acetamido residue of the GalNAc is replaced by a cyclic amide fused to the C1 and C2 carbons of the galactose ring.



14
R_1= H, CH_2OH

15
R_2= H (Ser), CH_3 (Thr)

Figure 4.4.2 Structures of conformational constrained C-glycoside analogues

In 2016, Mellet, Fernandez, Peregrina *et al.*[83] reported sp^2-iminosugar analogues of TnThr and TnSer determinants. The O6 atom was substituted with part of a cyclic carbamate, which limits the chair conformation flexibility (**15** - **Figure 4.4.2**). The analogues **15** were used as building block in peptide synthesis and APD[T*/S*]RP glycopeptides were obtained. Binding affinity against SM3 were measured and the best affinity has been shown for the analogue **15** linked to a Thr residue, although lower than the natural derivative.

4.5 Thr vs. Ser

A crucial issue, underlying theme of the last two chapters is whether the Ser or Thr residue linked to the GalNAc portion are equivalent in terms of immunogenicity of the resulting Tn antigen. This is indeed essential to focus the research, on the best determinants.

In the last few years, several papers have been published on the role of the first amino acid of the MUC1 peptide backbone and on the differences

between the 3D conformation of α-GalNAc-Ser and α-GalNAc-Thr.[84,85] It has clearly been shown that α-GalNAc-Thr is more rigid with respect to α-GalNAc-Ser: in TnThr the pyranose ring is oriented almost perpendicular to the peptide, whereas in TnSer the GalNAc portion is nearly parallel.[55,86–88] Live *et .al* ascribed this difference to steric repulsion between the sugar endocyclic oxygen and the β-methyl group of the Thr residue.[89] Instead, Corzana *et al.* ascribed the different orientation to the presence of water pockets formed by the pyranose ring with the peptide, which involves different atoms in TnThr and TnSer.[55] Regardless of the reason, such a difference is crucial in the recognition processes and seems to explain the higher binding affinity of TnThr for anti-Tn (Tn218), anti-MUC1 (SM3 *e.g.*) antibodies and several lectins.[87,88,90,91] The structural insights into the α-GalNAc-Thr motif and its higher specificity represent the rationale for the development of a *locked* TnThr mimetic.

CHAPTER

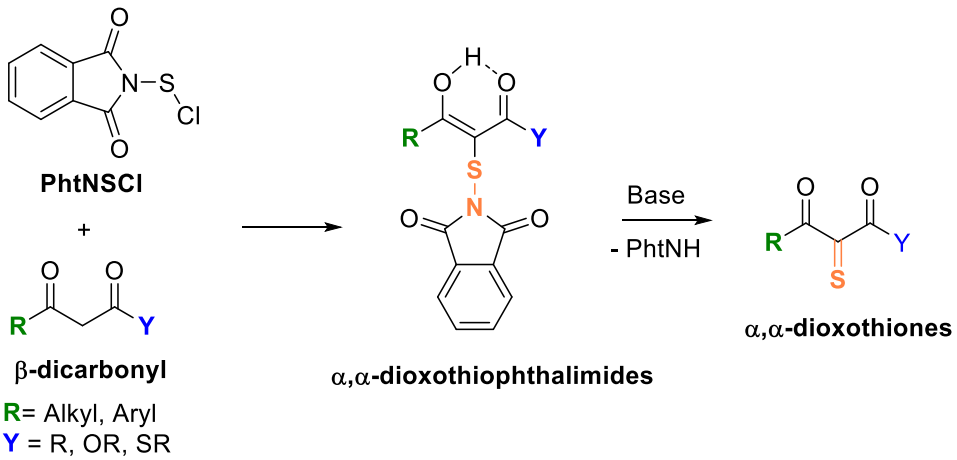
5

A locked α-TnThr mimetic

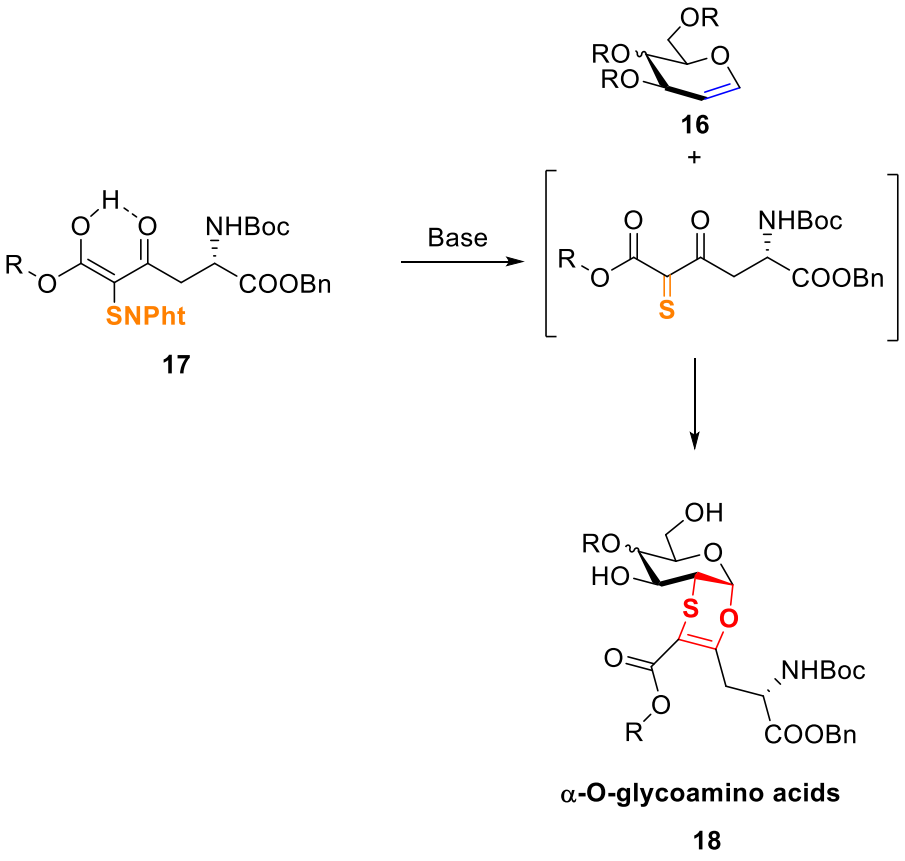
5.1 A powerful inverse electron-demand [4+2] cycloaddition

In the 90's, Capozzi, Nativi, Menichetti *et al.* reported that α,α'-dioxothiones (**Scheme 5.1.1**), generated from α,α'-dioxothiophthalimides (prepared by reaction of β-dicarbonyl compounds with phthalimidesulfenyl chloride as key reagent), are efficient electron-poor dienophiles.[92,93]

From this evidence, in 2004, Nativi *et al.* described the synthesis of 2-deoxy-2-thio-α-O-glycohomoglutamates (**Scheme 5.1.2**) by chemo-, regio-, and stereoselective cycloadditions between electron-rich glycals **16** and electron-poor aspartic acid derivatives **17.**[81] This novel chemical approach was developed for an easy and totally stereoselective preparation of α-O-glycoamino acids **18**.



Scheme 5.1.1 Generation of α,α'-dioxothiones



Scheme 5.1.2 Synthesis of 2-deoxy-2-thio-α-O-glycohomoglutamates **18**

In 2009, this powerful Diels-alder reaction was applied for the synthesis of the structurally rigid mimetic of the α-TnThr antigen, **13** (**Figure 5.1.1**).[82] The galactosyl moiety of the **13** is "*frozen*" in the $^4C_1$ conformation, the native α-O glycosidic linkage is preserved and the Thr residue is mimed by the cyclic ene-amide. In addition, a carboxylic group, suitable huck for the insertion of further portions, characterizes the structure of **13**. It is worthy of noting, the improved stability to chemical and enzymatical hydrolysis of the unnatural α-O glycosidic linkage featuring this mimetic.

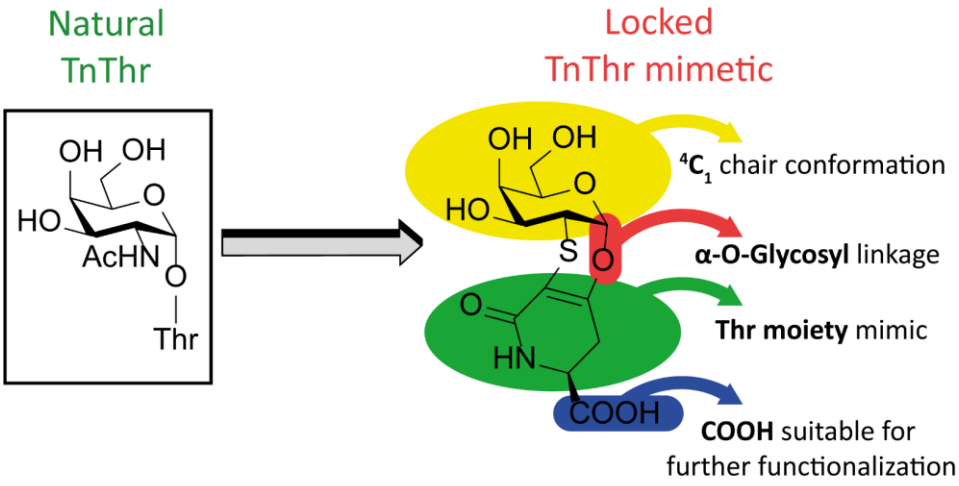


Figure 5.1.1 Structures of the native Tn and of the *locked* TnThr mimetic **13**

Nuclear magnetic resonance (NMR) and molecular modelling (MM) conformational studies, run by Jìmenez-Barbero, Nativi *et al.*[82] of **13** compared to the natural antigen confirmed the retain of the correct conformation and the recognition by *viscumin album* agglutinin (VAA), a lectin which selectively binds galactose residues. Furthermore, the binding properties of three other lectins, the plant lectin *Erythrina cristagalli* (ECL), the human macrophage C-type lectin (MGL), which is an

endogenous Tn receptor, and the snail protein *Helix pomatia agglutinin* (HPA), were evaluated by saturation transfer difference (STD) NMR studies.[94] ECL showed similar recognition of **13**, and the native ligand α-OMe-GalNAc, competing for the same binding pocket with a binding preference for **13**. In contrast, the STD-derived epitope mapping for **13** binding to MGL and ECL were different from that found for the natural substrates.

Altogether, the data reported suggest that the structural constrain characterizing **13** does not impair the pharmacophoric presentation of the glycosyl moiety.

5.2 Multivalent presentation of the TnThr mimetic 13

A milestone in the development of carbohydrate-based vaccines is represented by the discovery of the so-called *cluster glycoside effect*[95], that can be explain as " *an enhancement in valence-corrected binding activities of multivalent saccharide ligands*".[96] Even if the physical basis of this effect are not complete clear,[97] an improvement in activity of multivalent glycoconjugates compared to the corresponding monovalent ligand is observed. From this evidence, the synthesis of *multivalent glycoconjugates* of the TnThr mimetic **13**, to trigger an efficient immune response, has been widely explored by *Nativi et al.*[98–100] Among the plethora of natural and synthetic scaffolds available: *a)* nanoparticles of different nature, properties, and potential applications and *b)* a synthetic cyclopeptide carrier, were selected for the multivalent presentation of **13.**

5.2.1 Superparamagnetic iron oxide nanoparticles

A multivalent glycoconjugate nanosystem, composed of superparamagnetic iron oxide nanoparticles (MNPs) decorated with the TnThr mimetic **13** (GMNPs - **Figure 5.2.1.1**) was initially employed.[98] Superparamagnetic iron oxide nanoparticles have largely been exploited for supporting and carrying biomolecules, like antibodies, drugs or antigens, allowing to combine the main features of an inorganic magnetic core with a bioactive organic coating. Moreover, the heat they can release by the application of an external magnetic alternating field can be exploited in the treatment of tumors also by magnetic fluid hyperthermia (MFH).[101]

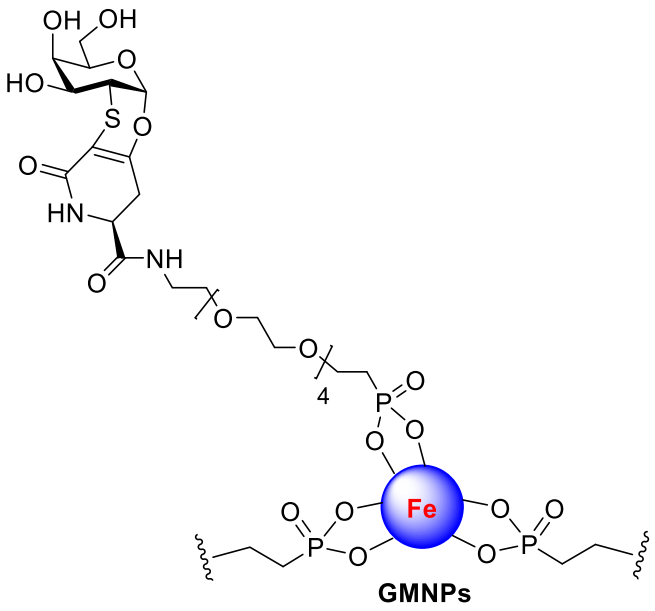


Figure 5.2.1.1 Glycosyl superparamagnetic iron oxide nanoparticles (GMNPs)

The absence of toxicity and endotoxin contamination was demonstrated and the effects of GMNPs with respect to the monovalent **13**, on mouse

monocyte/macrophage quiescent cell line (RAW 264.7) as model of human macrophages, were evaluated. Macrophages treated with **13**, and GMNPs, clearly showed the ability to endocyte only the nanosystem and not the monovalent antigen. Noteworthy, GMNPs have been able to activate macrophage effector functions, inducing gene expression and protein release of TNF-α at levels comparable to those obtained with the *golden standard* LPS.[98]

5.2.2 Dextran-based single chain polymer nanoparticles

In the quest for more biocompatible nanosystems, due to the concerns affecting metal nanoparticles as *in vivo* delivery methods,[102] Marradi, Chiodo, Nativi *et al*. developed dextran-based single-chain polymer nanoparticles (DXT-SCPNs) grafted with the mimetic **13** as nanocarriers (TnThr-DXT-SCPN - **Figure 5.2.2.1**).[100]

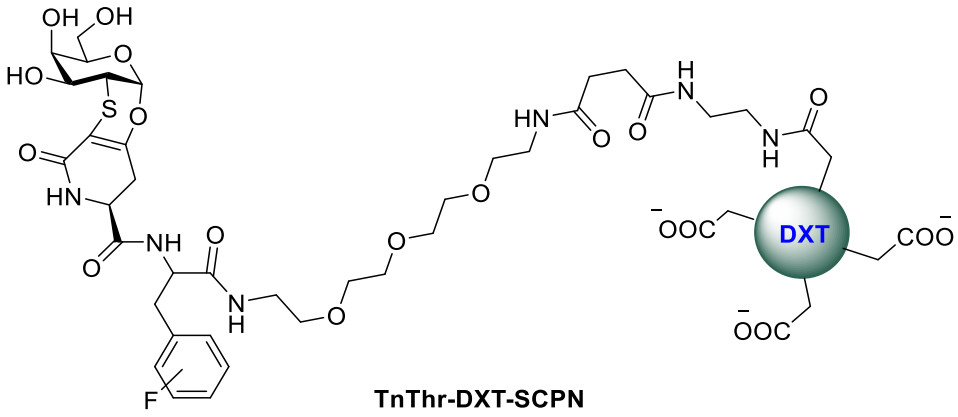


Figure 5.2.3.1 Schematic representation of TnThr-DXT-SCPN system

The stimulation of human peripheral blood monocytes (PBMC) was studied for the TnThr-DXT-SCPNs. The multivalent nanoparticles have been able to trigger the secretion of IL-6 and IL-10, similarly as the positive

control LPS and as already reported for mucin glycoproteins.[103] Non-glycosylated DXT-SCPNs induced nor IL-6 or IL-10 secretion.

5.2.3 A tetravalent synthetic cyclopeptide carrier - RAFT

In the design of a fully synthetic cancer vaccine candidate, the decapeptide scaffold, regioselectively addressable functionalized template (RAFT) developed by Dumy *et al.* (see **Chapter 4.3**) was decorated with four residues of the locked TnThr mimetic **13** (TnThr mime-RAFT - **Figure 5.2.3.1**).[99] The cyclopeptide carrier also displays an immunostimulant peptide epitope (OvaPADRE), linked to one of the Lys residue.

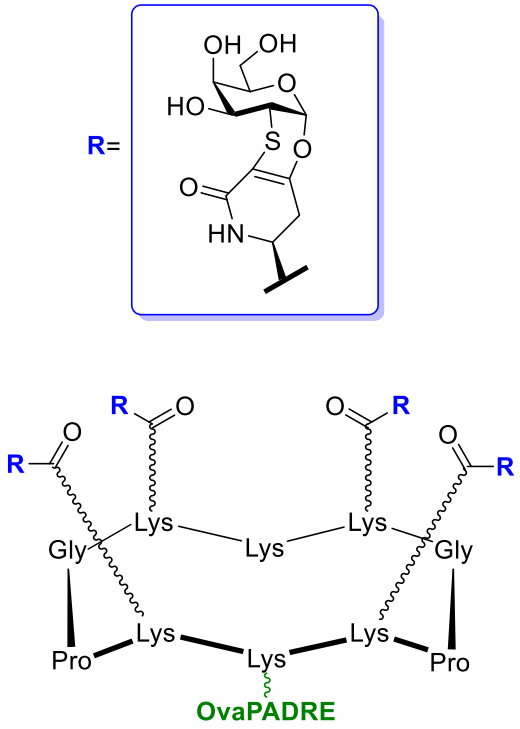


Figure 5.2.3.1 Structure of the TnThr mime-RAFT

The safety, immunogenicity and protective efficacy of TnThr mime-RAFT were successfully tested in a breast cancer animal model (BALB/c mice). This vaccine prototype induced a robust and long-lasting IgG/IgM antibody response, generated protection in mice reducing tumor progression, and increased survival through a B-cells-mediated mechanism.

CHAPTER

6

Aim of the project

Worldwide, cancer is a leading cause of death, and the research for new treatments to "*fight this struggle*" is an ever-changing matter. Active immunotherapy has proved to be an effective and valuable option, and tumor-associated antigens (TAs) emerged as potential therapeutic markers. TAs glycomimetics, that preserve the active conformation of native epitopes, represents a powerful "*tool*" to overcome low bioavailability and poor stability of the natural counterparts. In the last few years, mucin-related antigens, largely overexpressed in adenocarcinomas, have been widely studies and Nativi group reported on how a rigid mimetic of the TnThr antigen can be an extremely interesting scaffold in the development of stables antigen analogues.

In this framework, inspired by the excellent results obtained, **the aim of my PhD project** was to develop and investigate **the potential applications** of **structurally constrained mimetics of MUC1-related antigens**.

Part II

A TnThr mimetic

CHAPTER

7

A breast cancer vaccine candidate

7.1 Introduction

As widely discussed in the Introduction (**Chapter 2**), immunotherapy has become an established cornerstone of cancer treatment and therapeutic cancer vaccines are matter of vast research.[104] Cancer vaccines that display non-self-antigens to the immune system, may elicit a specific antitumor immune response leading to cancer cells' killing (**Chapter 2 - Section 2.3**).[105,106] In this context, cell surface-exposed carbohydrates, deserve a special attention, because several cancer cells can be differentiated from normal cells by the presentation on their surface of abnormal glycosylation motifs (Tumor Associated Carbohydrate Antigens, TACAs).[52] In fact, although poorly immunogenic and usually T cells independent, carbohydrate antigens, if properly presented to the immune system, can be recognized as non-self and induce an effective immune response.[107–110] Indeed, it has been reported about some cancer patients

able to induce natural auto-antibodies directed against native TACAs and with an improved survival rate.[111,112]

In the last two decades, a great attention has been caught by the MUC1 antigen, α-Tn (or Tn) (**Chapter 3**). Structurally simple, the Tn determinant has been the TACA of choice for the development of fully or semi-synthetic cancer vaccines.[113–116]

In the design of glycoconjugate vaccines, important issues shall be taken into account, in particular: *a)* due to the typical weak binding interactions between lectins and single glycans, a multivalent presentation of individual carbohydrate antigens linked to immunogenic carriers shall be assembled to increase the binding interaction and trigger a robust recognition event,[95,97] *b)* to elicit TACA-specific IgG antibodies, vaccine constructs should include Toll-like receptors (TLR)[117–119] or T helper peptides, as internal adjuvant.[114]

Among immunogenic carriers, CRM_{197} (Cross Reactive Material 197), a 58 kDa genetically detoxified mutant of Diphtheria Toxin (DT), obtained by a single Glycine to Glutamic amino acid substitution at position 52 (G52E)[120] is currently one of the most effective carrier proteins. It is licensed to treat bacterial infections and largely found in human conjugate vaccines.[121,122] An enhancement in immunogenicity can indeed be achieved with CRM_{197} glycoconjugate vaccines as result of the glycoconjugate's processing and presentation of carrier-derived T cell epitopes by MHC II molecules to T_h CD4+ cells. These cells can thus produce cytokines that stimulate hapten-specific B cell clones enhancing a carbohydrate-specific plasma response (*hapten-carrier effect* - **Chapter 1** - **Section 1.3.4**).[123,124]

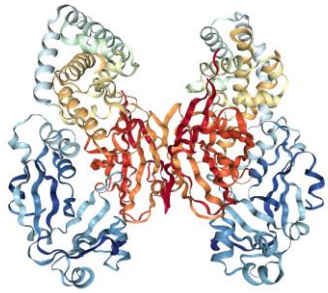


Figure 7.1.1 Crystal structure of CRM_{197} (PDB ID: 4AE0)

In breast cancer (BC) a correlation among the high level of Tn antigen expression in the primary tumor, tumor size and a poor prognosis is known, since 20 years.[39,125] Among the different BC kinds, the triple negative BC (TNBC) is the most difficult to treat for it is very aggressive, prone to metastasize and resistant to the current anti-BC therapies.[126]

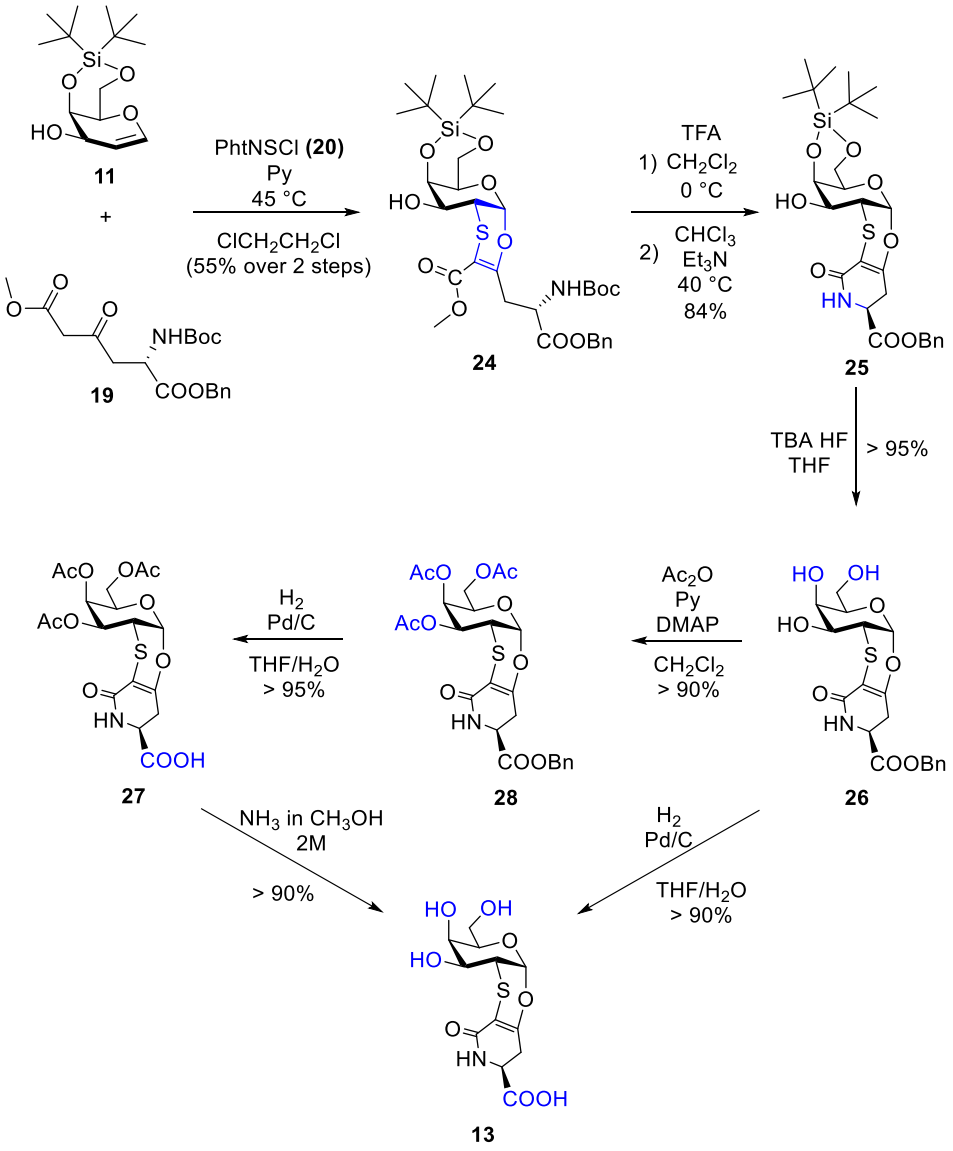
In this regard, taking advantage from the stable and antigenic TnThr mimetic recently developed [82] (**Chapter 5**), in this Chapter are reported an optimized synthesis of the "locked" TnThr mimetic **13** and the characterization and immunological evaluation of CRM_{197} glycoconjugates presenting residues of **13**, as candidate vaccine to treat non responsive TNBC.

7.2 Results and Discussion

7.2.1 An optimized synthesis of the TnThr mimetic

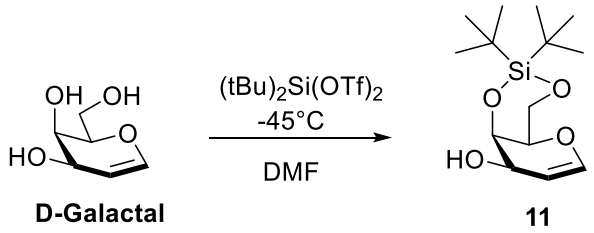
A gram scale synthesis of **13**, firstly reported in 2009,[82] has been accomplished and herein reported (**Scheme 7.2.1.1**). The essential building blocks to achieve the saccharide scaffold **13** are the protected D-

galactal **11**, the β-ketoester **19**, derived from aspartic acid and the phthalimidesulfenyl chloride (PhtNSCl) **20**.



Scheme 7.2.1.1 Large scale synthesis of the TnThr mimetic **13** and of its acetylated derivatives **27** and **28**

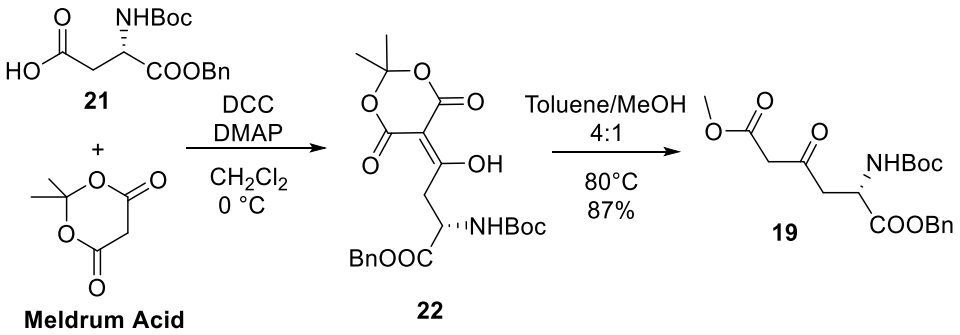
Compound **11** can be easily obtained by reacting D-galactal (**Scheme 7.2.1.2**) with *di-tert-butylsilyl bistriflate* [$(tBu)_2Si(OTf)_2$] in N,N-dimethylformamide (DMF) as solvent. The simultaneous protection of the hydroxyl groups in positions four and six of galactal is essential to make the bottom (α) face of the electron-rich dienophile the less hindered, thus the preferred one, in the cycloaddition reaction with the electron-poor diene **12**. The crude product **11** was used without purification in the following cycloaddition.



Scheme 7.2.1.2 Synthesis of the electron-rich dienophile **11**

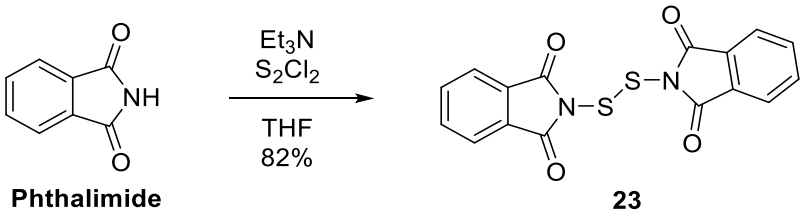
The synthesis of the β-ketoester **19** was optimized (**Scheme 7.2.1.3**) using the *N,N'-dicyclohexylcarbodiimide* (DCC) as coupling reagent in the presence of *4-(dimethylamino)pyridine* (DMAP), instead of *isopropyl chloroformate* (IPCF)[81] to activate the carboxylic residue of the protected aspartic acid, this allowed to increase the yield (from 70 to 87%) reducing costs. Thus, aspartic acid benzyl ester, suitably protected as N-Boc (*tert*-butyloxycarbonyl - **21**), was transformed in the presence of Meldrum acid into derivative **22**. By heating the crude **22** in a 4:1 mixture of toluene/CH_3OH, the β-ketoester **19** was obtained, in an 87% yield.

The key reagent PhtNSCl was formed from phthalimide which was firstly transformed (**Scheme 7.2.1.4**) into the *N-N*'- dithiobisphtalimide



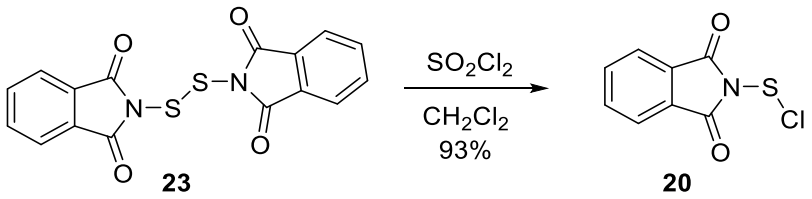
Scheme 7.2.1.3 Synthesis of the β-ketoester **19**

derivative **23**. White needles were obtained by recrystallization from $CHCl_3/CH_3OH$ (2:1) (82% yield).



Scheme 7.2.1.4 Synthesis of the *N*-*N*'- dithiobisphtalimide derivative **23**

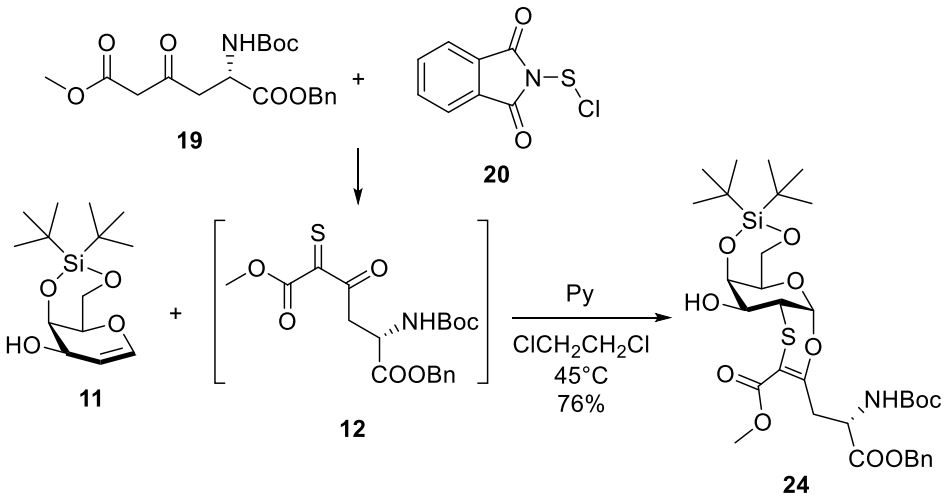
Disulfide **23**, in the presence of sulfuryl chloride (SO_2Cl_2) in tetrahydrofuran (THF) as solvent, formed the desired phthalimidesulfenyl chloride **20** in a 93% yield with a purity of 94% (by NMR) (**Scheme 7.2.1.5**).



Scheme 7.2.1.5 Synthesis of the phthalimidesulfenyl chloride **20**

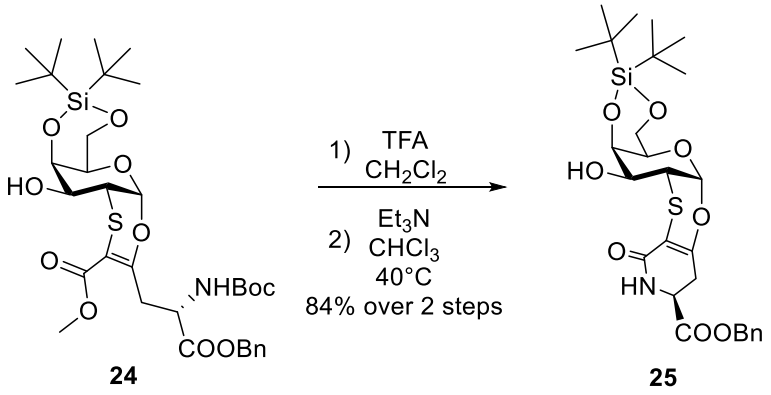
The α-α'-dioxothione **12** was generated *in situ* by reacting **19** with **20**, under mild basic conditions (**Scheme 7.2.1.6**) and trapped by reaction

with crude **11**. After flash chromatography purification, the cycloadduct **24** was isolated in a 55% yield (over two steps) as single diastereomer.



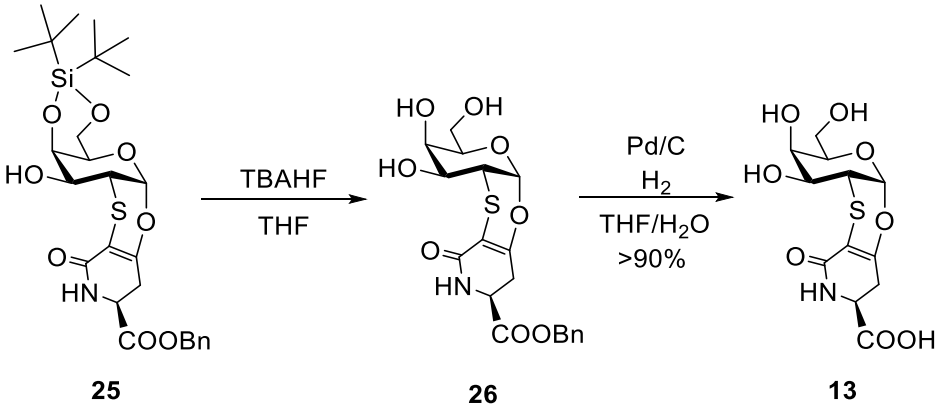
Scheme 7.2.1.6 *In situ* generation of the α-α'-dioxothione **12** and synthesis of cycloadduct **24**.

Removal of the Boc protecting group from **24** with trifluoroacetic acid (TFA), and treatment with triethylamine (Et_3N) (**Scheme 7.2.1.7**), lead to the formation of the lactam **25** in an 84% yield.



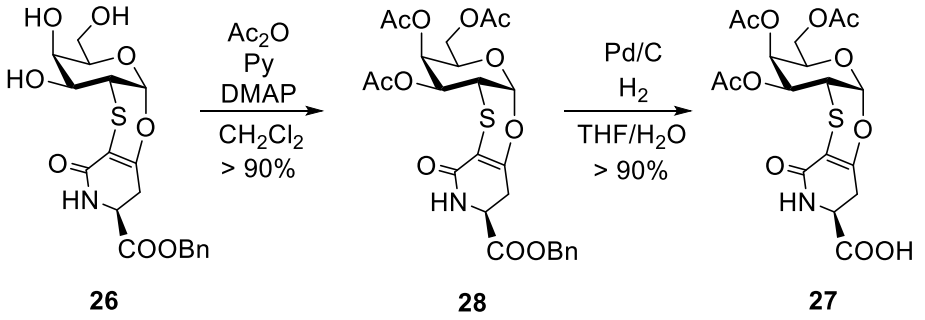
Scheme 7.2.1.7 Synthesis of the lactam **25**

Derivative **13** can be easily obtained from **25** (**Scheme 7.2.1.8**), by removal of the silylidene protecting group with freshly prepared tetrabutylammonium hydrogen fluoride (TBAHF) in THF to form **26** which was hydrogenated with palladium as catalyst. The unprotected mimetic **13** was thus formed as a white solid, in a high yield (> 90%).



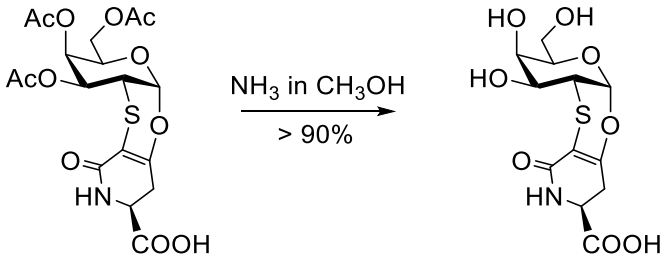
Scheme 7.2.1.8 Synthesis of the TnThr mimetic **13**

Another interesting precursor that can be easily synthesize from **26** is the peracetylated derivative **27**, presenting the free carboxylic acid (**Scheme 7.2.1.9**).



Scheme 7.2.1.9 Synthesis of the acetyl derivative **27**

The reaction of **26** in the presence of acetic anhydride (Ac_2O), pyridine (Py) and a catalytic amount of DMAP gave **28** that, after a palladium catalyzed hydrogenation, was converted into the acetyl derivative **27** (> 90% yield). Noteworthy, **27** can also be converted into the unprotected **13** after treatment with a solution of NH_3 in CH_3OH (2M) (**Scheme 7.2.1.10**).



27 **13**

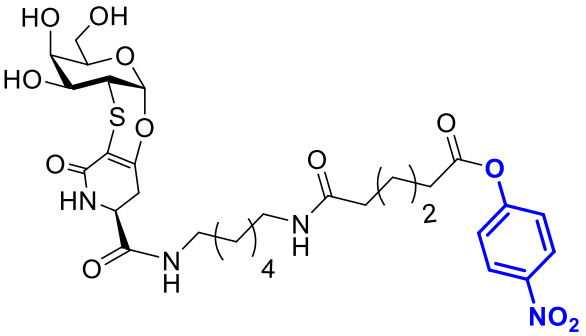
Scheme 7.2.1.10 Deacetylation of **27** to fully unprotected **13**

In the efficient synthesis of the TnThr mimetic **13** described herein some building blocks that we will find recurrently are reported. In particular, **25** is the *glycosyl acceptor* which will be used for the synthesis of the TFThr mimetic (**Chapter 9**), and compound **26** is the precursor of the glycosyl acceptor which will be used for the synthesis of the STnThr mimetic (**Chapter 10**).

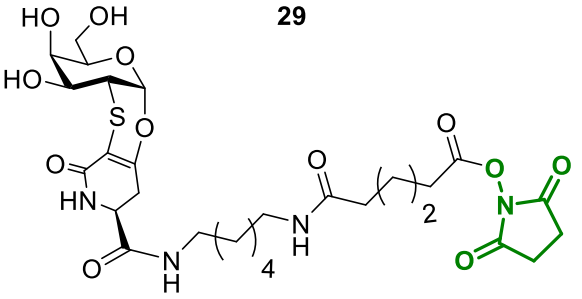
As already said, the presence on **27** of the free carboxylic residue, makes it suitable for the introduction of different linkers by efficient reactions like an amide bond formation.

7.2.2 Synthesis of TnThr mimetic-CRM_{197} glycoconjugates

To decorate the clinically validated carrier-adjuvant protein CRM_{197}, differently activated derivatives of **13** were synthesized (**29** and **30** - **Figure 7.2.2.1**)



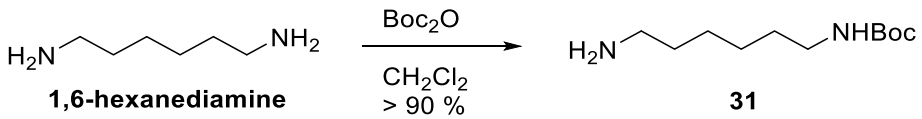
29



30

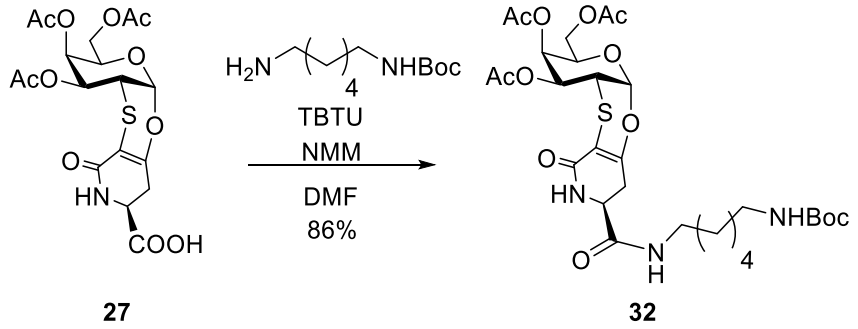
Figure 7.2.2.1 Structure of activated derivatives **29** and **30**

To obtain the desired linker for the mimetic **27**, we firstly synthesized the mono-Boc protected amine **31**. A straightforward synthesis (**Scheme 7.2.2.1**) of the N-Boc-1,6-hexanediamine **31** was achieved from 1,6-diaminohexane and di-tert-butyl dicarbonate (Boc_2O) in dichloromethane as solvent.



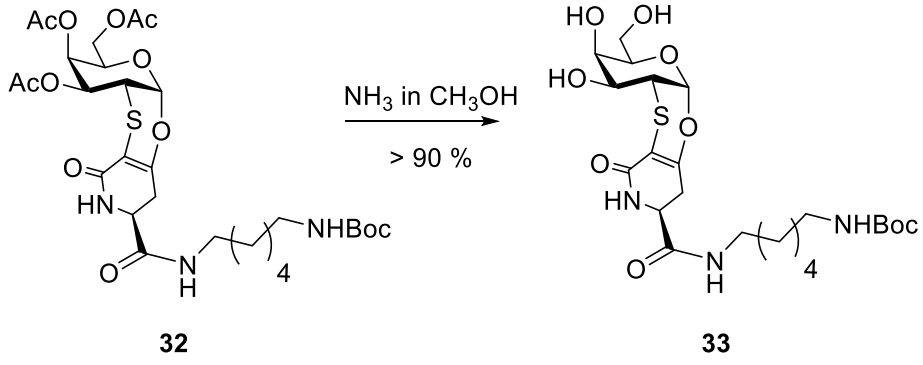
Scheme 7.2.2.1 Synthesis of N-Boc-1,6-hexanediamine **31**

Isolated **31** was then reacted with the acetyl derivative **27** (**Scheme 7.2.2.2**), in the presence of 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate (TBTU) and N-methylmorpholine (NMM), in dry DMF as solvent, affording derivative **32** in an 85% yield.



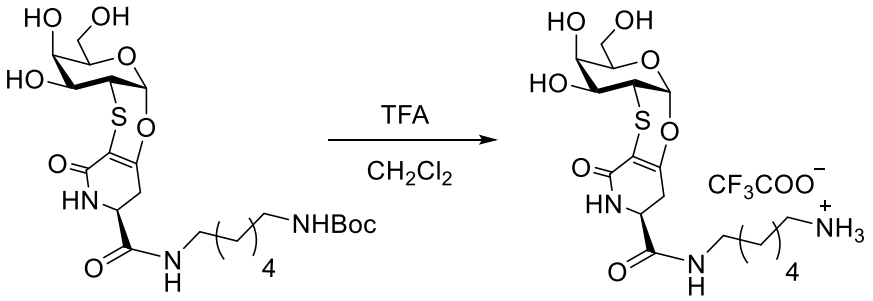
Scheme 7.2.2.2 Synthesis of derivative **32**

The acetyl protecting groups were then removed by treating **32** with a solution of NH_3 in CH_3OH (2M) to give **33** (> 90%) (**Scheme 7.2.2.3**).



Scheme 7.2.2.3 Deacetylation and synthesis of **33**

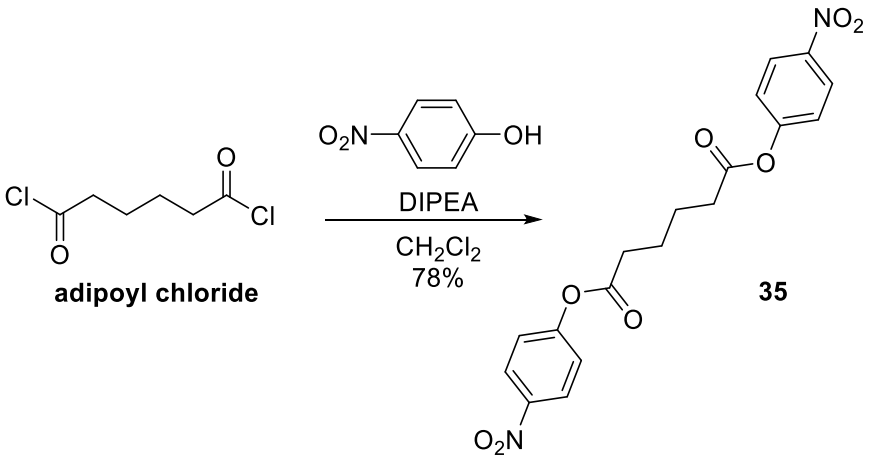
The Boc protecting group was cleaved with TFA and the trifluoroacetic salt **34** obtained was isolated and used without any further purification (quantitative) (**Scheme 7.2.2.4**).



33 **34**

Scheme 7.2.2.4 Synthesis of the trifluoroacetic acid salt **34**

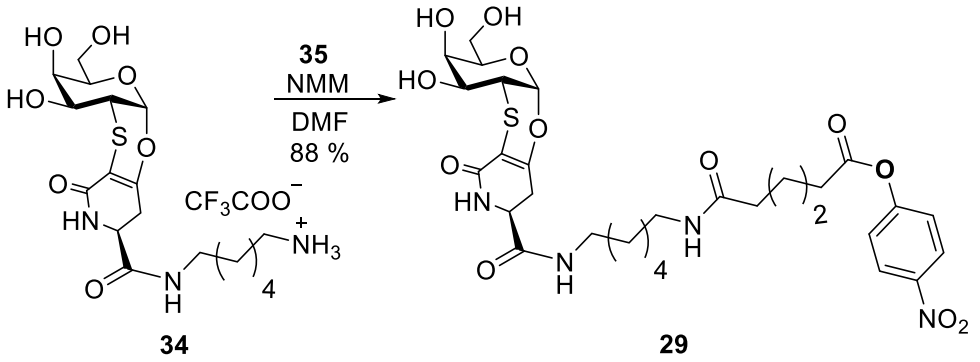
The bifunctional activated para-nitrophenyl derivative **35** was then synthesized (**Scheme 7.2.2.5**) from 4-nitrophenol and adipoyl chloride in the presence of N,N-diisopropylethylamine (DIPEA) in a 78% yield.



adipoyl chloride **35**

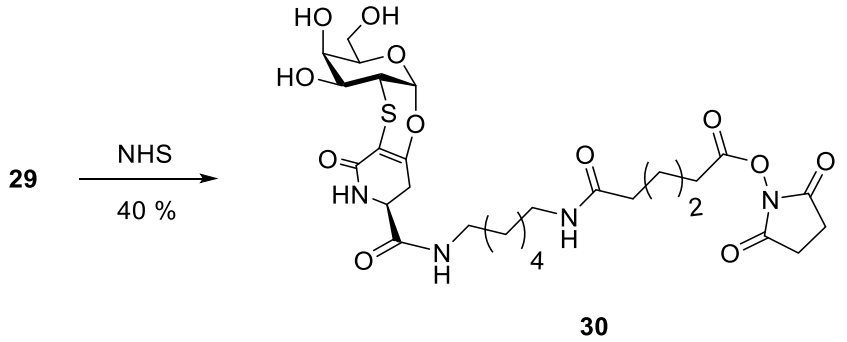
Scheme 7.2.2.5 Synthesis of the bifunctional activated para-nitrophenyl linker **35**

The linker **35** was then reacted with crude **34** and NMM in DMF (**Scheme 7.2.2.6**) to afford, after flash chromatography, compound **29** in an 88% yield, over two steps.



Scheme 7.2.2.6 Synthesis of the activated derivative **29**

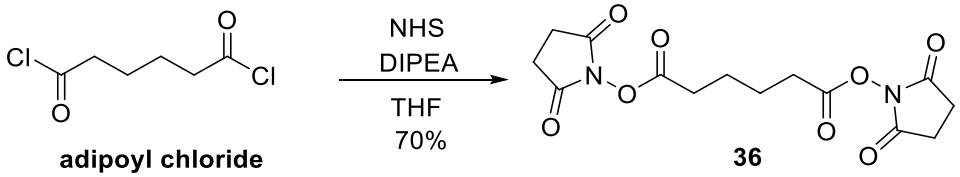
Replacing the para-nitrophenyl group with an N-hydroxy succinimidyl residue, derivative **30** was firstly synthesized by reacting **29** (**Scheme 7.2.2.7**) with of N-hydroxysuccinimide (NHS), to give **30** in a 40% yield.



Scheme 7.2.2.7 Synthesis of the activated derivative **30**.

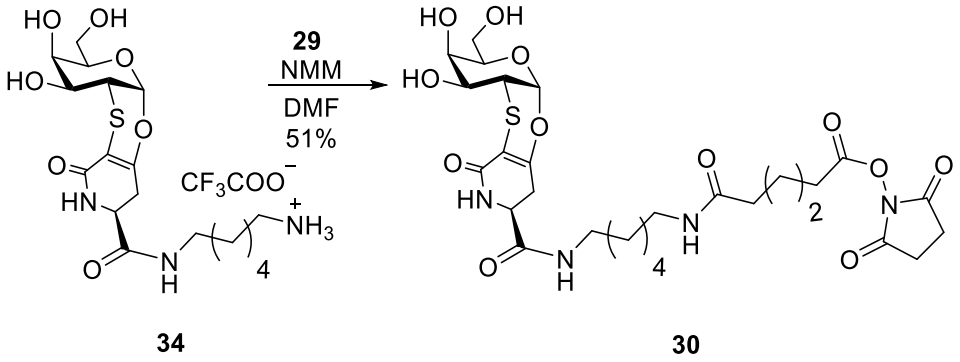
More efficiently, the bis-succinimide derivative **36** was synthesized as reported for **35** (**Scheme 7.2.2.8**). After crystallization from acetone (70%

yield), **36** was used for the synthesis (**Scheme 7.2.2.9**) of **30** by reaction with **34** in the presence of NMM and in DMF as solvent.



Scheme 7.2.2.8 Synthesis of bis-NHS linker, **36**

Compound **30** was isolated in a 51% yield, over two steps. The lower yield compared to the synthesis of **29** can likely be ascribed to a higher reactivity of the NHS-group. The hydrolysis of the activated NHS ester was indeed observed, especially during the purification step.

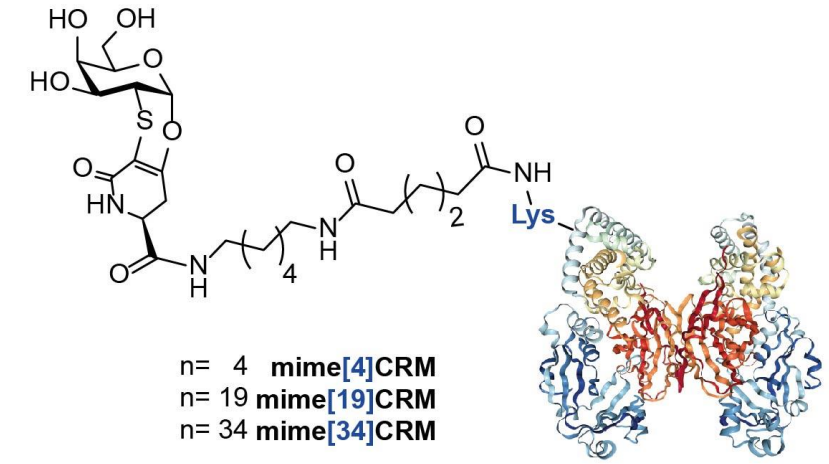


Scheme 7.2.2.9 Synthesis of **30** from **36**

CRM_{197} was then treated with **29** or **30** and as expected, a different loading of the protein was yielded. (**Figure 7.2.2.2**) Upon conjugation with **29**, up to 4 synthetic antigens were coupled to CRM_{197} (**mime[4]CRM**) while conjugation with the more soluble and reactive ester **30**, allowed to vary the loading of the protein upon the amount of antigen used, and up to 19 or 34 synthetic glycans were grafted to CRM_{197} to form the

glycoconjugates **mime[19]CRM** or **mime[34]CRM** respectively. In order to have a control for the planned biological tests, the glycoconjugate (Glc-CRM) was also prepared. Activated D-glucose (Glc) was synthesized according to published conditions.[127]

The number of linked glycosides for each glycoconjugate was estimated by 2,4,6-trinitrobenzene sulfonic acid (TNBS) assay and confirmed by Matrix-Assisted Laser Desorption (MALDI) spectra.



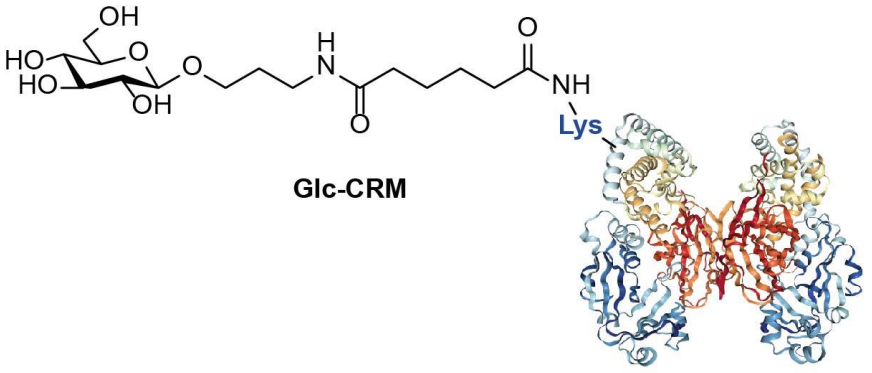


Figure 7.2.2.2 Representation of CRM_{197} glycosylated with n= 4, 19 or 34 residues of TnThr mimetic or with Glc residues.

7.2.3 Structural characterization of CRM_{197} glycoconjugates

1D ^{1}H-NMR spectra of non-functionalized CRM_{197} (**Figure 7.2.3.1, A**), **mime[4]CRM** (**Figure 7.2.3.1, B**) and **mime[19]CRM** (**Figure 7.2.3.1, C**) were recorded.

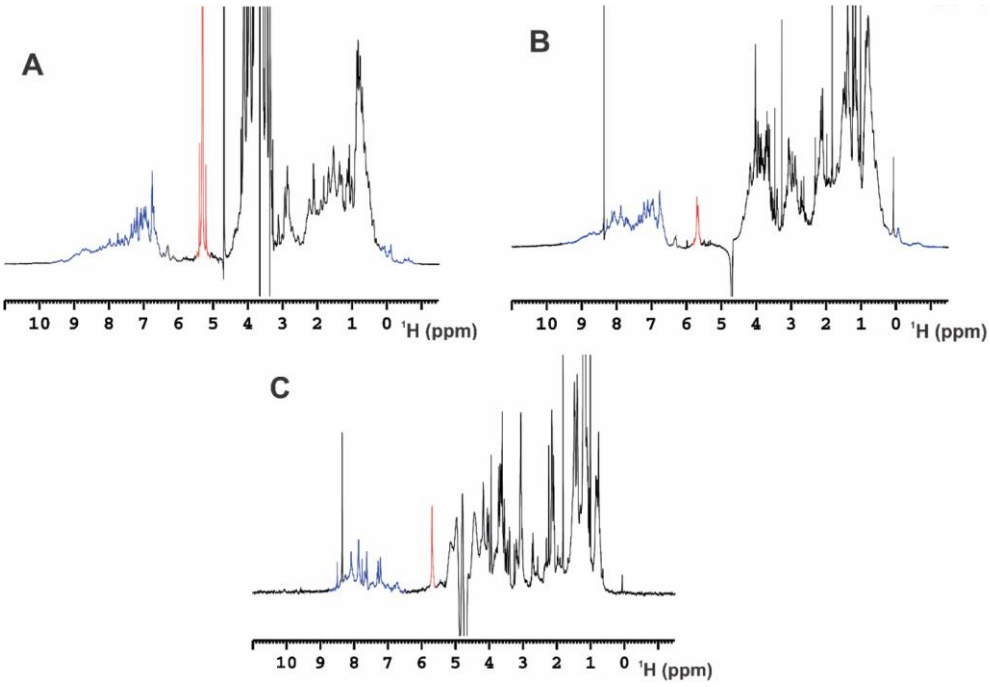


Figure 7.2.3.1 1D ^{1}H NMR spectra at 900 MHz and 298 K of CRM_{197} (**A**), **mime[4]CRM** (**B**) and **mime[19]CRM** (**C**). The signal of the anomeric proton of sucrose present in the formulation buffer of free CRM_{197} (**A**) and the anomeric protons of the synthetic glycans conjugated to the protein (**B** and **C**) are highlighted in red. The signals of the amide and methyl protons are highlighted in blue. In the free CRM_{197} (**A**) and in **mime[4]CRM** (**B**) the amide signals are spread out over 3 ppm (from 9.5 to 6.5 ppm) while in **mime[19]CRM** (**C**) the spreading is sizably reduced (from 8.5 to 6.5 ppm) and the signals of methyl protons in the region from 1 to -1 ppm are completely missing.

Noteworthy, **mime[4]CRM** preserves its original folded structure, while the protein conjugated with 19 residues is largely unfolded. The dispersion of the NMR signals in the regions of the amide protons and methyl protons provided the main indicators for the folding state of proteins. In the spectrum of **mime[4]CRM**, these resonances are well separated, with some signals in the region between 0.5 and -1 ppm, whereas in **mime[19]CRM** the signals have narrow chemical shift dispersion without resonance lines of the protein below 0.5 ppm. Indeed, small-angle X-ray scattering (SAXS) (**Figure 7.2.3.2)** spectra showed a clear aggregation in the case of **mime[19]CRM** but globular and compact particles in the case of CRM_{197}, confirming that this glycoconjugate is aggregated in solution.

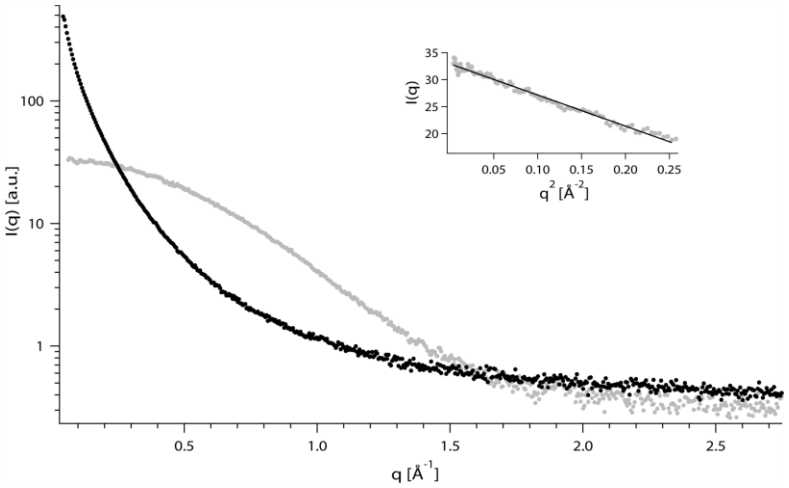


Figure 7.2.3.2 SAXS spectra of **mime[19]CRM** (black) and CRM_{197} (grey), recorded at the BioSAXS beamline BM29 (ESRF, Grenoble, FR) at 37 °C.

Molecular dynamic (MD) studies performed on mimetic **13** showed clear evidence of the conformational restriction imposed by the additional rings [root-mean-square deviation (RMSD) (heavy atoms) = 0.84 Å, **Figure 7.2.3.3**] and highlighted the orientation of the amino acid with respect to

the sugar moiety relative to natural TnThr. Being aware that a protruding of the TnThr mimetic residues linked to the protein is required for any immune recognition, as an example, we performed also 0.5 μs MD simulations on **mime[4]CRM** in explicit water (**Figure 7.2.3.4**). According to these simulations: *a)* the four **13** were conjugated to the most exposed lysine residues *b)* the tertiary structure of CRM_{197} is not significantly altered upon the chemical modifications and *c)* all TnThr mimetic residues are exposed to the solvent, essential condition for the efficacy of the vaccine.

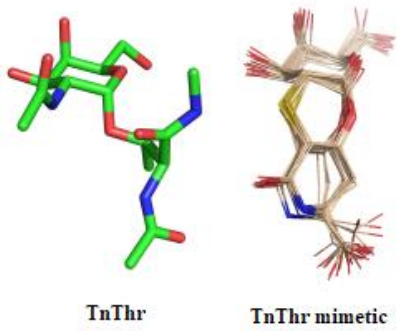


Figure 7.2.3.3 Representative conformer of natural **TnThr** (as a diamide derivative) in water solution, together with the structural ensemble derived from 0.5 μs MD simulation of mimetic **13**.

Although the protein folding is not an essential requirement and can even be positive exposing the T-cell epitopes, need of a proper control (folded CRM_{197}) to discern from any non-specific immunological effect was considered. Results reported in the present section determined the choice of **mime[4]CRM** for subsequent *in vitro* and *in vivo* studies.

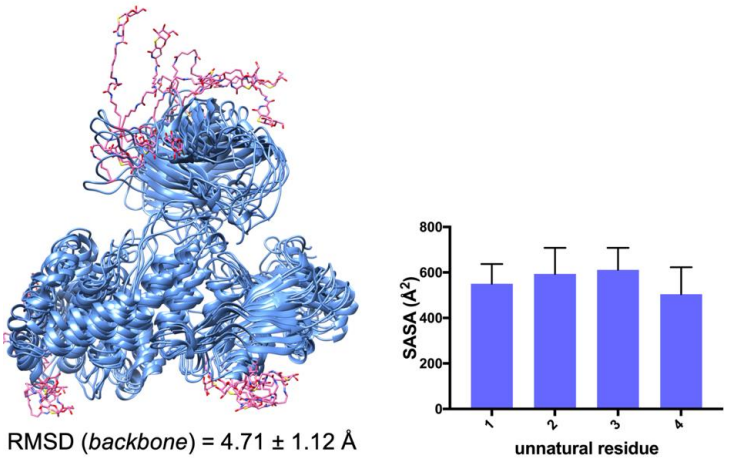


Figure 7.2.3.4 Representative snapshot derived from 0.5 μs MD simulations of CRM_{197} (PDB ID CRM197: 4AE0) upon chemical modification with 4 molecules of mimetic **13** together with the RMSD of the backbone of CRM_{197} and the solvent accessible surface area (SASA) for the unnatural residues. The protein is shown as ribbons and the unnatural residues are represented as sticks.

7.2.4 Binding of mime[4]CRM to anti-Tn antibody

In order to confirm: *a)* the accessibility of **13** to the solvent (*vide infra*), *b)* the preservation of the correct conformation of the natural Tn *c)* the improvement in immunogenicity of the multivalent presentation of **13**, the affinity of monovalent **13** and **mime[4]CRM** to the anti-Tn monoclonal antibody Tn218[128,129] was assessed by surface plasmon resonance (SPR) (**Figure 7.2.4.1**).

While the isolated mimetic **13** (**Figure 7.2.4.1 - B**) exhibited an affinity comparable to that of natural TnThr, with a $K_D = 1.60 \times 10^{-2}$ M and 1.25×10^{-2} M respectively, the multivalent construct **mime[4]CRM** showed a $K_D = 1.15 \times 10^{-5}$ M (**Figure 7.2.4.1 - A**). This result proved the accessibility of **13**

to the antibody and showed a binding improvement by presenting multiple copies of the TnThr mimetic on the surface of the protein paving the way for subsequent *in vitro* ed *in vivo* test.

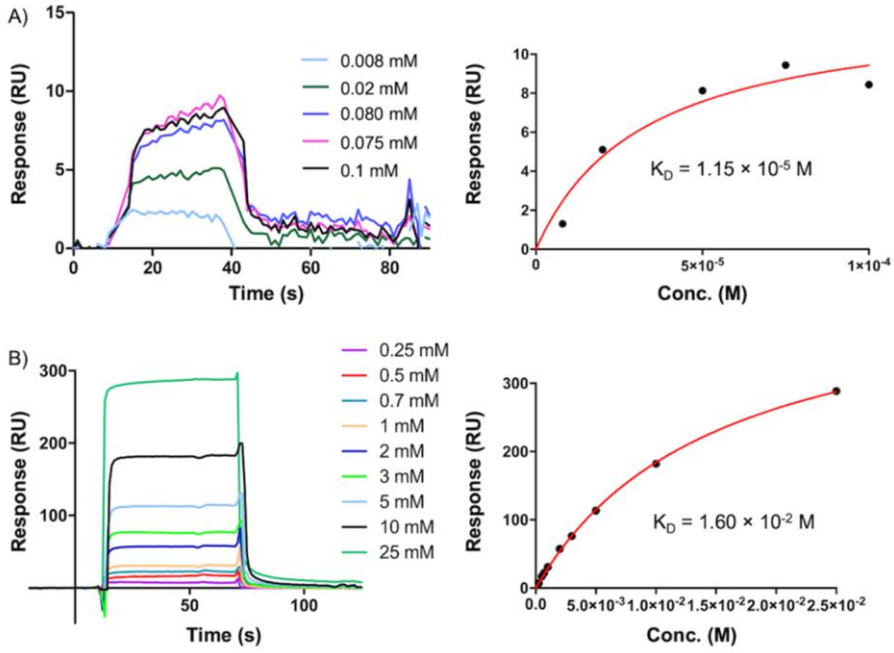


Figure 7.2.4.1. Representative SPR curves and fit obtained for **mime[4]CRM** (**A**) and isolated mimetic **13** (**B**) towards monoclonal antibody Tn218.

7.2.5 *In vitro:* activation of human dendritic cells (DCs)

DCs are central regulators of the adaptive immune response (**Chapter 1 - Section 1.3**) and important actors in elicitation of anti-tumoral responses.

The potential role of **mime[4]CRM** in DCs' activation and/or maturation, to check the possible immunogenicity of this new vaccine candidate in humans, was investigated. DCs were differentiated from peripheral blood adherent mononuclear cells (PBMC) of healthy donors and the expression of markers CD80, CD83 and CD86 after stimulation for 48 h with

mime[4]CRM or controls was checked (**Figure 7.2.5.1**). Flow cytometry analysis showed that **mime[4]CRM** induced activation and maturation of DCs as demonstrated by the increased expression of CD83 and CD86 markers. In contrast, the controls Glc-CRM and native TnThr peptide did not.

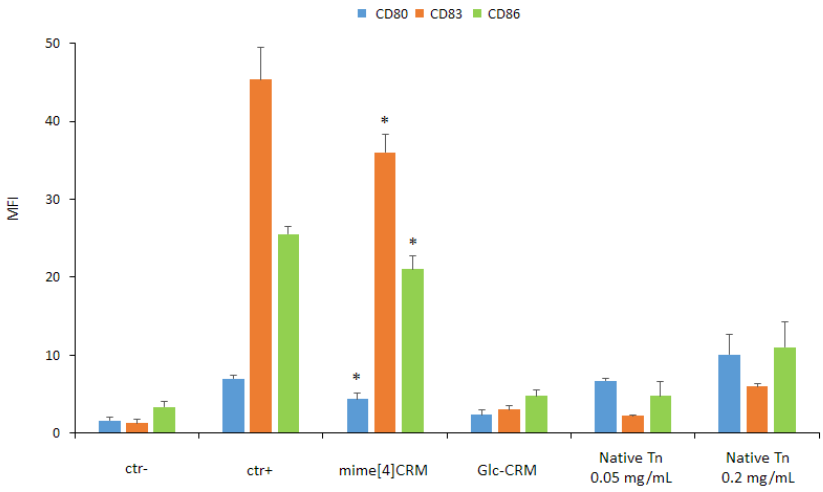


Figure 7.2.5.1 Evaluation of Human DCs activation: DCs isolated from three healthy donors were stimulated with negative control (ctr-), positive control (ctr+, LPS), **mime[4]CRM**, Glc-CRM and native AlaProAspThr(Tn)ArgPro. The difference between **mime[4]CRM** treatment and the ctr- was assessed using paired t test; *P < 0.05.

7.2.6 *In vivo:* evaluation of the therapeutic effect of mime[4]CRM by using TNBC transplanted model

Capitalizing on the unique structural features and interesting *in vitro* properties of **mime[4]CRM**, its potential anti-tumorigenic activity against the challenging TNBCs was evaluated. First, the overall lack of toxicity on:

a) 4T1-Luc cell line, as a cellular model of highly metastatic TNBC,[130] *b)* human PBMC and *c)* human TNBC cell line (MDA-MB231) exposed to **mime[4]CRM** (from 2.3 to 57.5 µg/mL) or controls (**Glc-CRM** and CRM_{197} alone) was verified. No influence on viability and growth rate on cells populations was observed upon treatment with **mime[4]CRM** (**Figure 7.2.6.1**).

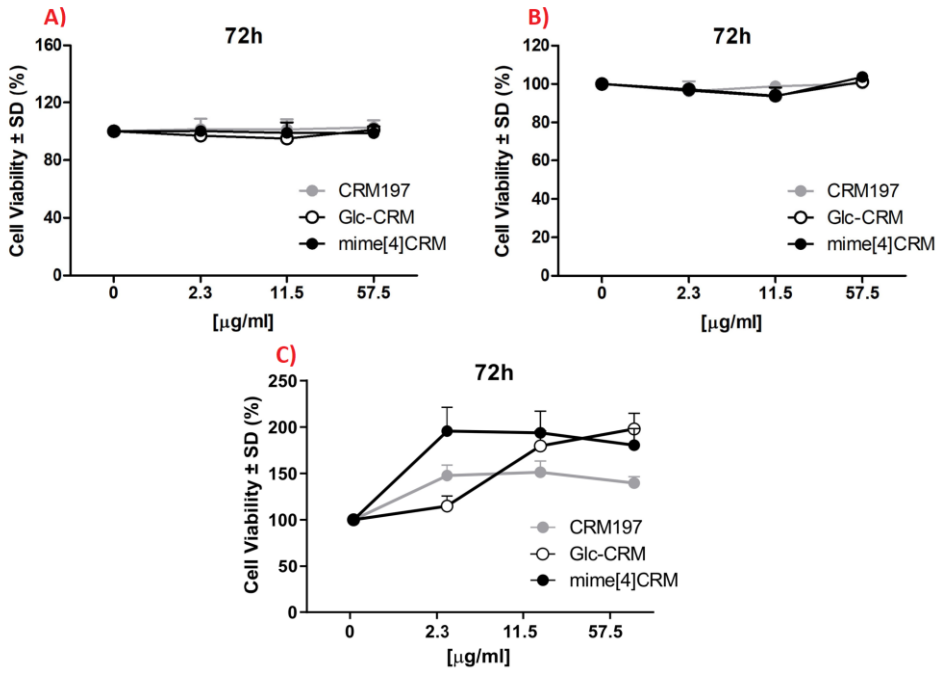


Figure 7.2.6.1 Cell viability upon exposure to **mime[4]CRM** 4T1 mouse breast cancer cells (**A**), Human PBMC (**B**) and MDA-MB231 human TNBC (**C**).

The potential anti-tumorigenic action of **mime[4]CRM**, together with the activation of immune cells within the tumor microenvironment, was then evaluated in vivo by performing a preclinical study. For this purpose, Tn expressing murine 4T1-luc cells (stably expressing Firefly Luciferase

gene)[131] were implanted into the mammary fat pad gland of immunocompetent syngeneic mouse model (BALB/c mice).

Noteworthy, the use of immunocompetent mice properly allows to test the efficacy of immunomodulating compounds. Moreover, recent data showed that wild-type and huMUC1-transgenic mice produce equivalent antitumoral response against a native Tn-containing candidate vaccine.[132]

After that tumors were established, mice were imaged (Bioluminescence Imaging, BLI) at the time of implantation (day 0, T0) then subcutaneous administered with **mime[4]CRM** (n=9) or with CRM_{197} as vehicle for control group (n=10) weekly for 6 weeks (**Figure 7.2.6.2**).

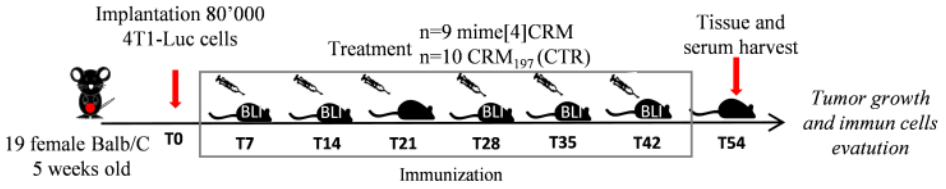


Figure 7.2.6.2 Scheme of the *in vivo* preclinical trial

Mice were imaged every 7 days to monitor tumor growth *in vivo* (**Figure 7.2.6.3**) and then sacrificed at T28, T42 and T54.

The results obtained clearly showed that the trend of total flux was significantly reduced after a week (T14 - *P < 0.01) and also after three weeks (T28 - **P < 0.05) of treatments in the group of mice that received **mime[4]CRM** compared to the vehicle group (**Figure 7.2.6.4 - A**), thus demonstrating an anti-tumorigenic action *in vivo* of **mime[4]CRM**.

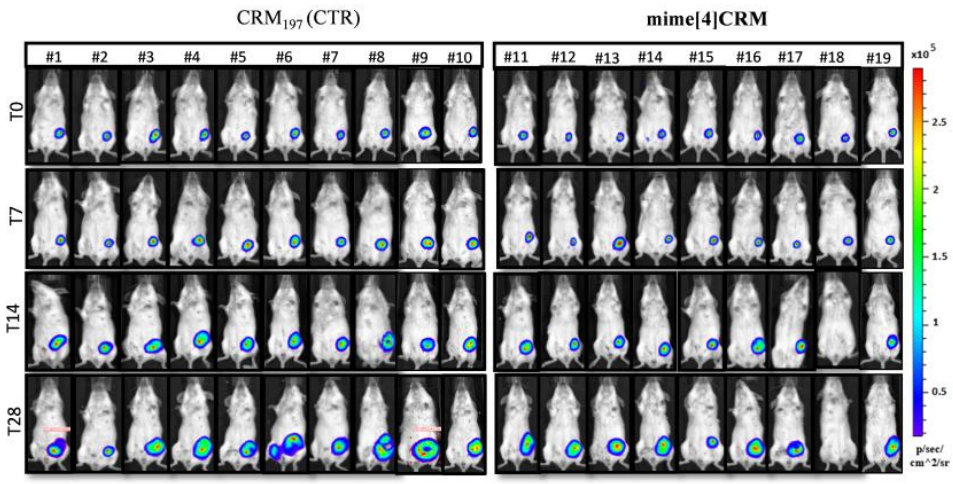


Figure 7.2.6.3 BLI images of mice orthotopically transplanted using 4T1-Luc cells

The absence of body and organs' weight loss (**Figure 7.2.6.4 - B**) in the treated mice confirmed that there was no significant acute toxicity induced by **mime[4]CRM**, in line with *in vitro* data.

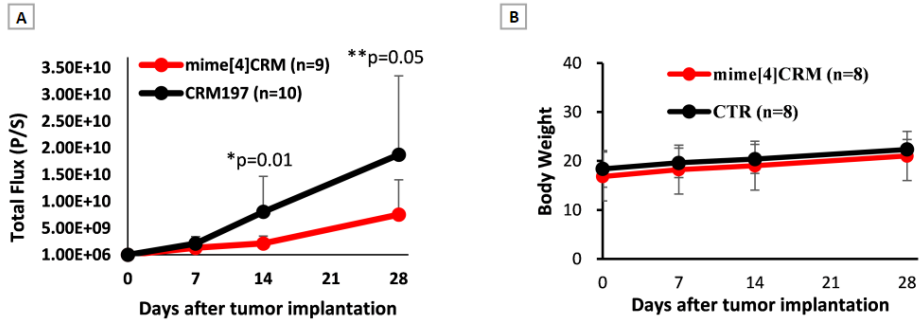


Figure 7.2.6.4 A) Quantification of photon emission from the region of interest (ROI) in mice treated with **mime[4]CRM** or with CRM_{197} **B)** Evaluation of drug toxicity using mice body weight in **mime[4]CRM** treated and control mice. The values are expressed as mean ±SD.

Immunofluorescence (IF) analysis of the lymphoid and myeloid cell populations infiltrating 4T1 mammary tumor sections from two mice (#9

and #19) at T28, treated with CRM_{197} and **mime[4]CRM** was performed by using antibodies directed against CD4 (T lymphocytes), TCR (T lymphocytes), CD11c (DCs), PD1 (immunosuppressive marker mostly expressed on T cells surface), FOXP3 (T_{REG} cells), CD68 (granulocytes), F4/80 (macrophages) and CD163 (pro-tumorigenic M2-macrophages) (**Figure 7.2.6.5**). These markers allowed to investigate the role of **mime[4]CRM** in triggering *in vivo* immune cells exerting inhibitory effects or promoting tumor spread. The data showed (**Figure 7.2.6.5**) increased levels of both T cells (CD4+, TCR+) and peripheral dendritic cells expressing CD4+ (CD4+, CD11c+) in mice treated with **mime[4]CRM**. This suggests that the candidate vaccine can modulate the recruitment of both DCs and T_h CD4+ cells in the tumor microenvironment. Furthermore, the tumor sections from the same treated mouse also showed a significant reduction of CD4+ T cells expressing the immunosuppressive marker PD1 (PD1+, CD4+) and the transcriptional factor FOXP3 (FOXP3+, CD4+), thus indicating a reduction in the quote of immunosuppressive T_{REG} cells. Data also showed a significantly reduction in the pro-tumorigenic M2-polarized tumor associated macrophages (TAMs; CD68+, CD163; F4/80, CD163+). Altogether, these IF staining indicates that **mime[4]CRM** vaccine has a clear potential role in modulating the recruitment and phenotype of CD4+ T cells, and, therefore on the polarization status of TAMs in the TNBC tumor microenvironment.

Furthermore, we investigated the potential ability of **mime[4]CRM** to decrease Prune-1 protein levels in treated mice. This protein had previously been reported to be overexpressed in 4T1 cells[133] and

positively correlated to advanced stage in metastatic BC.[134,135] However, immunofluorescence analyses performed on primary tumor sections derived from **mime[4]CRM**-treated and vehicle-mice did not show differences in Prune-1 levels, thus suggesting that the ability of the vaccine to reduce tumor outgrowth does not involve the down-regulation of Prune-1 levels *in vivo*, confirming its action on the modulation of immune infiltrating cells by activating the lymphocytes component.

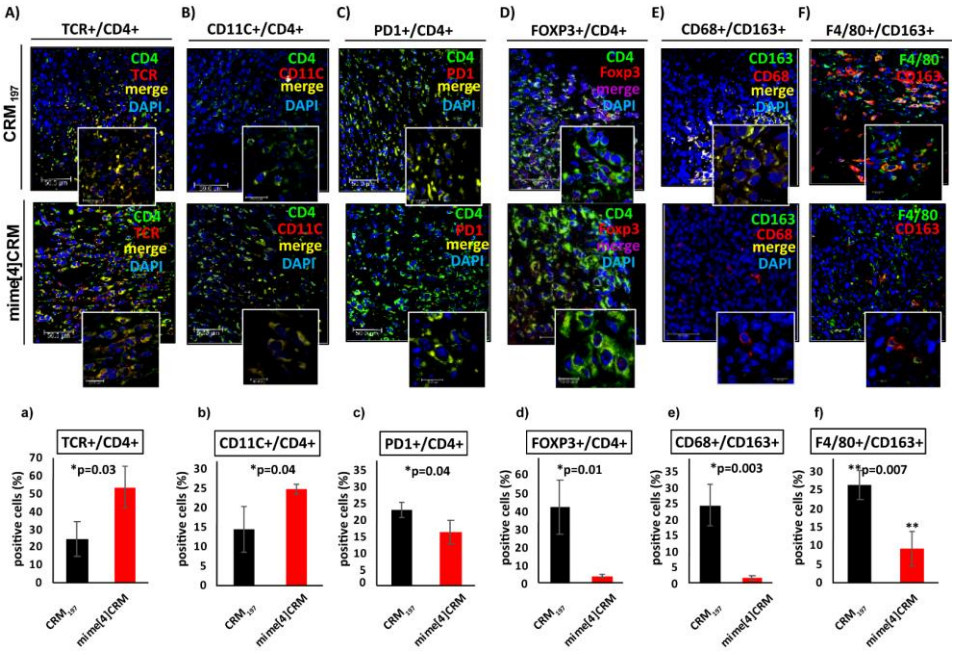


Figure 7.2.6.5 Immunofluorescence staining of primary mammary tumor sections from transplanted mice treated with **mime[4]CRM** or with CRM_{197} as a vehicle.

As 4T1 cells, used as a cellular model of TNBC, are highly metastatic, the presence of lung metastases was evaluated by performing *in vivo* and *ex-vivo* analyses. At T42, in vivo BLI analyses indicated the presence of lung metastases in seven of nine controls and four of eight **mime[4]CRM-**

treated mice (**Figure 7.2.6.6 - A, B**). In contrast, as determined through ex vivo BLI imaging, the presence of metastatic foci was detected in eight of nine control mice (seven mice at T42 and one mouse at T54, **Figure 7.2.6.6 - B**), and in four of eight **mime[4]CRM**-treated mice (one mouse at T42 and three mice at T54, **Figure 7.2.6.6 - A**). Therefore, although we cannot exclude the presence of lung micro-metastases at T42 in the **mime[4]CRM**-treated group of mice because they might not be detectable with the in vivo BLI technology, data showed a clear effect of our candidate vaccine in modulating lung metastasis development.

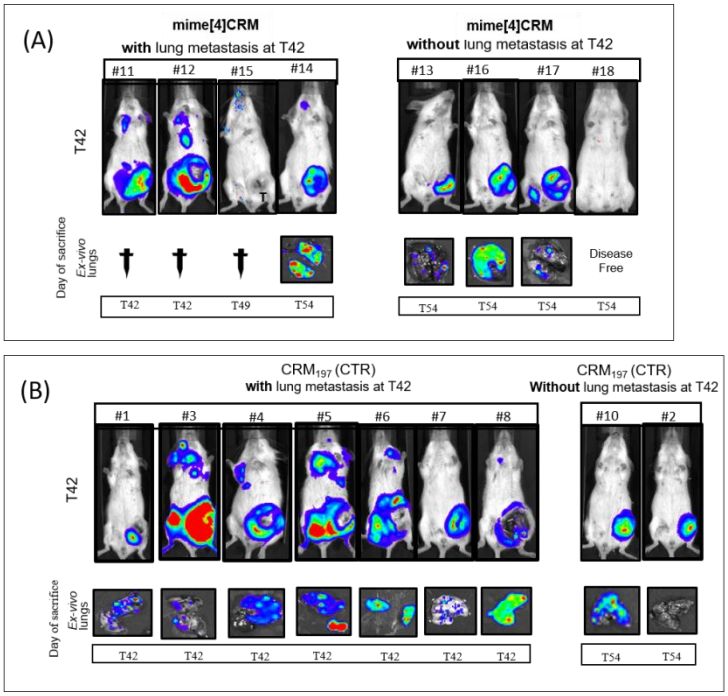


Figure 7.2.6.6 (A and B) *In vivo* BLI analyses at T42 (upper panel), and Ex vivo BLI analyses (lower panel).

Remarkably, cumulative survival analyses between the two groups of mice revealed 7 out of 9 (77.7%) controls and 2 of 8 treated (25%) mice died at

T42, thus indicating a clear trend of therapeutic benefit in **mime[4]CRM**-treated *vs.* control mice in the observed period.

7.2.7 Anti TnThr mimetic specific antibody response elicited by mime[4]CRM immunization

Results so far collected show that mice transplanted with 4T1 cancer cells and immunized with **mime[4]CRM** have a reduced tumor burden. The antibody response introduced by the vaccine was analyzed by measuring the IgG serum titers elicited upon immunization against the glycan portion of **mime[4]CRM**. To detect mimetic **13** specific antibodies, mice sera collected during immunization (T28 and T42) were tested by enzyme-linked immunosorbent assay (ELISA) on plates coated with the hexapeptide AlaProAspDAPArgPro **37** or with the glycosylated AlaProAspDAP(**13**)ArgPro hexapeptide **38** (**Figure 7.2.7.1**). The AlaProAspDAPArgPro peptide was selected as analogue of the AlaProAspThrArgPro (APDTRP) motif, minimum epitope of the VNRT region of hypoglycosylated TA(MUC1).[88] Immunization with **mime[4]CRM** induced predominantly IgG (almost exclusively IgG1, whereas IgG2a was barely detectable), and to a lower lever IgM (mean titer 1:1.350 and 1:600 respectively), compared with control mice (mean titer 1:600 and 1:100, respectively) (**Figure 7.2.7.2**).

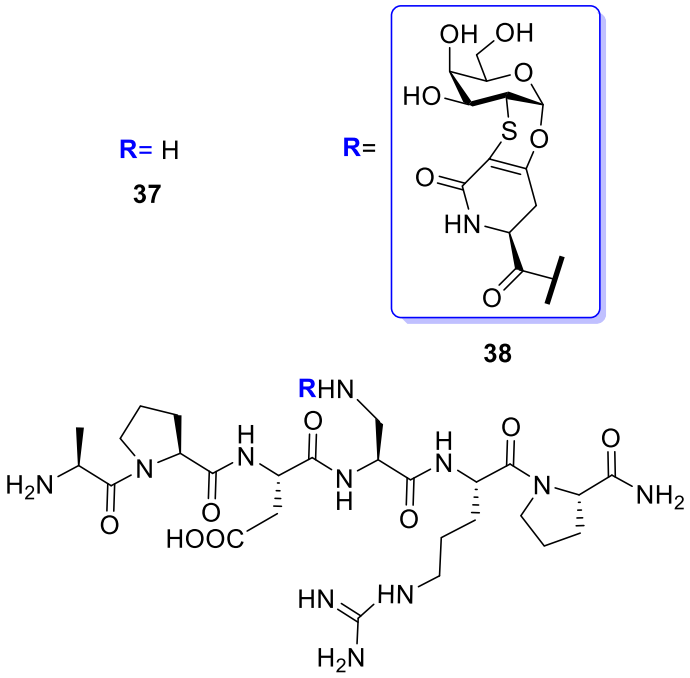


Figure 7.2.7.1 Structure of non-glycosylated **37** and TnThr mimetic-glycosylated **38** hexapeptides for ELISA assay.

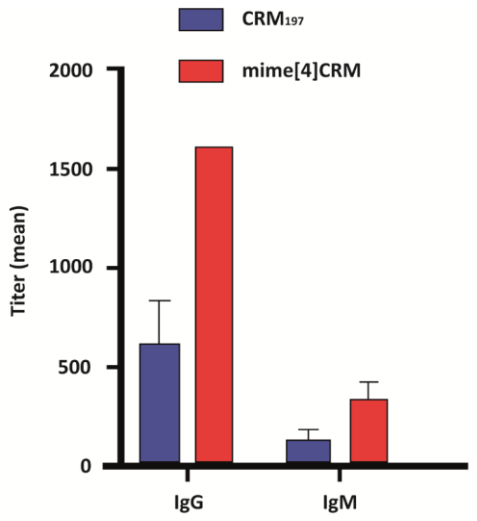


Figure 7.2.7.2 ELISA titers of IgM and IgG antibodies elicited against mimetic **13** by immunization with **mime[4]CRM** and control (CRM_{197}).

Sera were then tested for binding on the surface of 4T1 cancer cells by flow cytometry. Notably, mice immunized with **mime[4]CRM** elicited antibodies that can better recognize native Tn antigen expressed by BC cells than control mice with an approximately 3-fold increase in mean fluorescence intensity (MFI), (**Figure 7.2.7.3**).

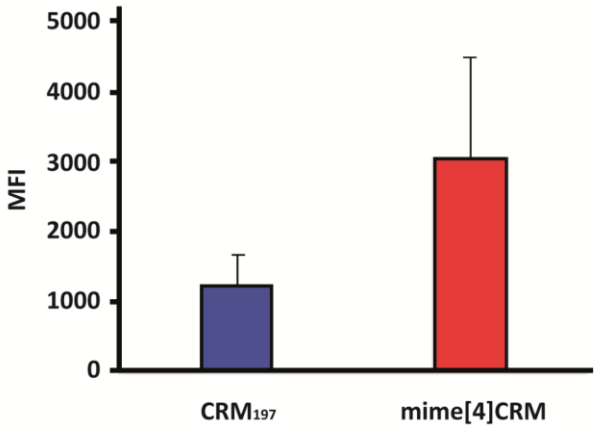


Figure 7.2.7.3 Average of MFI values of mice immunized with **mime[4]CRM** and control mice (CRM_{197}).

A preliminary analysis of serum cytokines revealed a higher level of anti-tumor Th17 cytokines after immunization with **mime[4]CRM** (**Figure 7.2.7.4**). This result was not unexpected, since the ability of Tn glycosylation in inducing IL-17 responses was recently described.[136] While the relevance of IL17+ T lymphocytes in anti-cancer immunity is still controversial, it may contribute to protective tumor immunity by recruiting effector cells to the tumor microenvironment.[137,138] In this study, the presence of IL-17 in serum of immunized mice is an additional evidence of **mime[4]CRM** immunomodulatory activity on T cells.

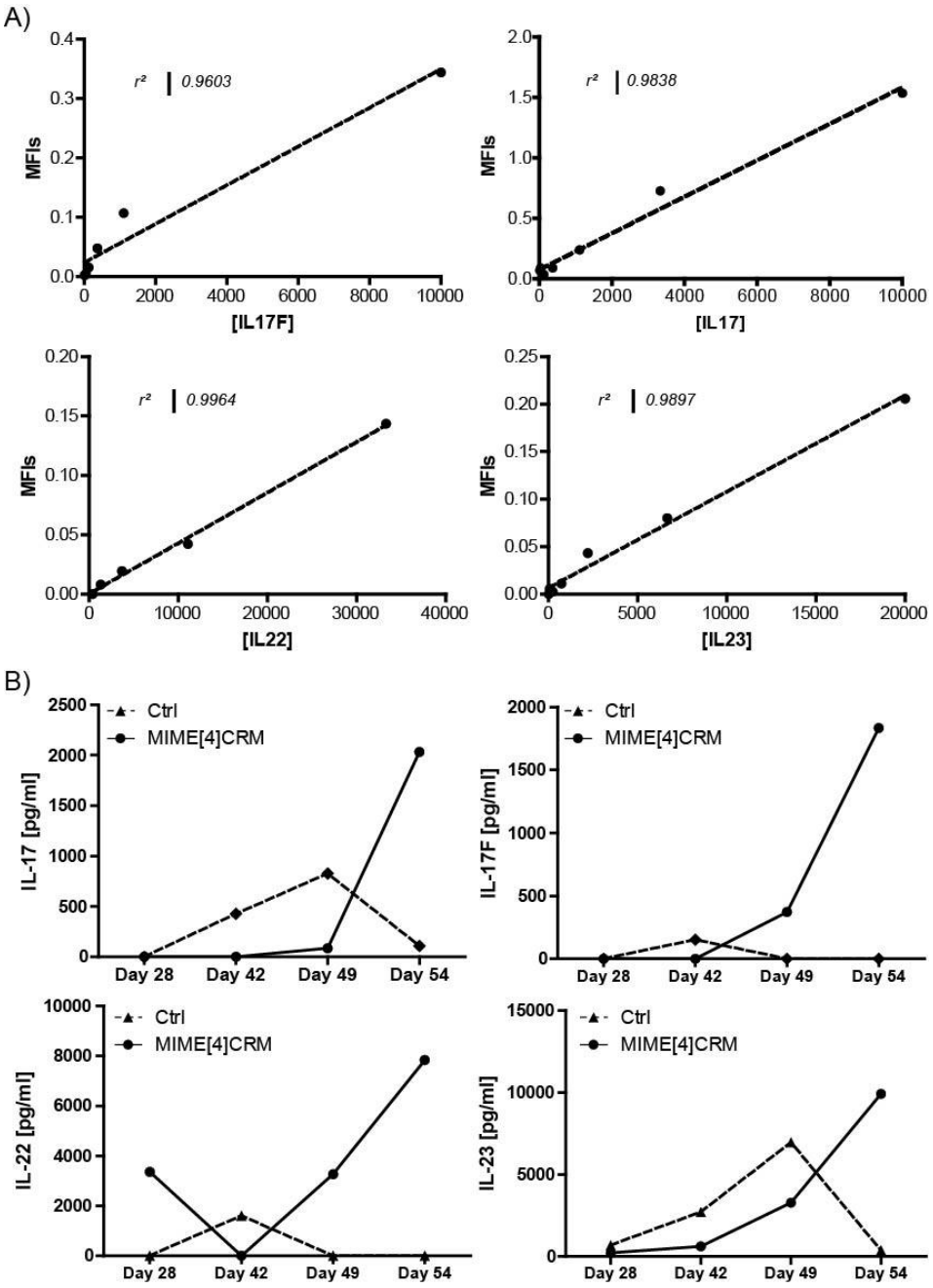


Figure 7.2.7.4 Quantification of serum cytokines showing a differential concentration in mice that developed or not metastasis after vaccination with **mime[4]CRM**

7.2.8 Isolation of mime[4]CRM-specific T cells.

In a preliminary studies twelve BC patients were enrolled to assess the presence of **mime[4]CRM**-specific T cells, in the peripheral blood and in the tumor-tissues. In detail, the PBMCs were cultured in the presence of medium alone or with **mime[4]CRM**. The presence of **mime[4]CRM**-specific T cells was documented in 8 (67%) out of the 12 patients. Thus, **mime[4]CRM**-specific T cells were isolated, by cloning the tumor infiltrating cells (TILs). A total of 90 T cell clones were obtained. and 15 (17%) were specific for **mime[4]CRM**. Of note, from an evaluation of cellular markers profile of the intra-tumoral **mime[4]CRM**-specific T cells, 69% were CD4+ while 31% were CD8+. This last evidence demonstrates the ability of **mime[4]CRM** to stimulate T cells in patients with BC. A compelling hypothesis is that this might be due to an activation of Tn-specific T cell response spontaneously elicited in cancer tissues.

7.3 Conclusions

In summary, an improved and refined synthesis of the *"locked"* TnThr mimetic **13** and its conjugation on the immunogenic carrier protein CRM_{197}, for the development of a novel vaccine candidate against the TNBC, was reported. As described, the protein loading can be easily modulated by choosing different activation methods and a complete structural characterization of glycoconjugates was achieved by combining NMR, SAXS and MD analyses. The accessibility of glycan epitopes to the solvent has been verified by SPR and the improved binding against an anti-Tn antibody (Tn218) confirms the potential increasing in immunogenicity

that can be achieved by a multivalent presentation. Of note, binding data underlined that the mimetic preserves the correct conformation of the natural determinant. Human DCs, properly activated *in vitro* by **mime[4]CRM**, showed the increased expression of maturation markers (CD83, CD86). Moreover, immunofluorescence staining showed that **mime[4]CRM** vaccine has a clear role in modulating the recruitment of DCs, T_h CD4+ cells, CD4+ T cells expressing PD1 and FOXP3 and pro-tumorigenic M2-polarized TAMs, thus resulting in impairing tumor progression *in vivo*. Furthermore, the presence of IL-17 in serum of immunized mice is an additional evidence of **mime[4]CRM** immunomodulatory activity on T cells. Relevantly and unprecedentedly for a TACA mimetic, the isolation of **mime[4]CRM**-specific TILs from BC patients highlighted the intrinsic immunogenicity of **mime[4]CRM** and the ability to stimulate a specific immune response. In conclusion, the result collected provide important evidence of **mime[4]CRM** fitness as new determinant for cancer vaccine in humans and they could be of great help for the rational design of structurally innovative cancer vaccines based on TACAs.

7.4 Experimental section

7.4.1 General methods

NMR of glycoconjugates: 1D ^{1}H NMR spectra of free CRM_{197} and of its glycoconjugates **mime[4]CRM** and **mime[19]CRM** were acquired at 298 K on a Bruker AVANCEIII-HD NMR spectrometer operating at 900 MHz, ^{1}H Larmor frequency, equipped with a cryogenically cooled probe. Free CRM_{197} in NaPi (10 mM) + sucrose (10% w/v) at pH 7.2 with 10% D_2O, while the glycoconjugates **mime[4]CRM** and **mime[19]CRM** were dissolved in NaPi (150 mM) at pH 7.5 with 10% D_2O.

SAXS recording method: SAXS data were recorded on the BioSAXS beamline BM29 at the European Synchrotron Radiation Facility (ESRF) Grenoble, France, using a 2D Pilatus detector at an X-ray wavelength of 1.008 Å. Data processing and reduction were performed using an automated standard ESRF beamline software (BSxCuBE)[139]and PRIMUS.[140] Concentration series (0.5, 1.0, 2.0, 5.0 mg/mL) were recorded for both samples and individual samples and buffers were exposed for 10 times 1s frames and grouped to improve statistics. For CRM_{197} the radius of gyration RG was extracted by the Guinier approximation[141] using PRIMUS.

Molecular dynamics (MD) simulations: The simulations were carried out with AMBER 18 package implemented with ff14SB[142] and GAFF[143] force fields. The structure of CRM_{197} was obtained from the protein data bank web site (PDB ID: 4AE0).The parameters and charges for the unnatural Tn and the corresponding linker were generated with the antechamber module of AMBER, using GAFF force field and AM1-BCC method for

charges.[144] The conjugated (or compound **1**) was immersed in a water box with a 10 Å buffer of TIP3P water molecules.[8] If required, the system was neutralized by adding explicit counter ions (Na^+). A two-stage geometry optimization approach was performed. The first stage minimizes only the positions of solvent molecules and the second stage is an unrestrained minimization of all the atoms . The system was then gently heated by incrementing the temperature from 0 to 300 K under a constant pressure of 1 atm and periodic boundary conditions. Harmonic restraints of 30 kcal*mol^{-1} were applied to the solute, and the Andersen temperature-coupling scheme was used to control and equalize the temperature. The time step was kept at 1 fs during the heating stages, allowing potential inhomogeneities to self-adjust. Long-range electrostatic effects were modelled using the particle-mesh- Ewald method.[145] An 8 Å cut-off was applied to Lennard-Jones interactions. Each system was equilibrated for 2 ns with a 2-fs time step at a constant volume and temperature of 300 K. Production trajectories were then run for additional 0.5 μs under the same simulation conditions.

Binding studies by Surface Plasmon Resonance (SPR): SPR experiments were performed with a Biacore X-100 apparatus (Biacore, GE) in HBS-EP buffer at pH 7.5 (Hepes 10 mM, NaCl 150 mM, ethylenediaminetetraacetic acid (EDTA) 3 mM, with 2% dimethyl sulfoxide (DMSO) and 0.005% Tween X100 as the running buffer at 25 °C. Tn218 antibody was immobilized on a CM5 sensor chip (Biacore, GE) following standard amine coupling method. Briefly, the carboxymethyl dextran surface of the flow cell 2 was activated with a 7-min injection of a 1:1 ratio of aqueous 0.4 M 1-ethyl-3-

(3-dimethylaminopropyl)carbodiimide (EDC) and 0.1 M sulfo-N-hydroxysuccinimide. Then, the antibody was coupled to the surface during a 7-min injection using several dilutions in 10 mM sodium acetate, pH 4.0. The unreacted active esters on the surface were quenched by a 7-min injection of aqueous 0.1 M ethanolamine-HCl (pH 8.0). The levels of immobilization were in the range of 3000-5000 resonance units (RU). Flow cell 1 treated as a flow cell 2 (amine coupling procedure) without protein was used as a reference. Prior to use, 50 mM stock solutions of **mime[4]CRM** were diluted to the final concentration in the running buffer. Typically, a series of different compounds was injected onto the sensor chip a flow rate of 30 µL/min for a period of 1min followed by a dissociation period of 1 min. No regeneration was needed. The concentrations used for affinity measurements were in the range of 0.05-1.5 mM. Sensogram data were double-referenced using the Biaevaluation X-100 software (Biacore, GE). The experimental data of affinity measurements were fitted to a one site-specific model binding using Prism software.

DCs markers' expression: DCs were differentiated from peripheral blood adherent mononuclear cells (PBMC) of healthy donors. PBMC were obtained by cell separation kit Magnetic Cell Separation (MACS - Miltenyi Biotec, Bergisch Gladbach - Germany). The cells were suspended and cultured in 6-well culture plates (2x106/mL) in RPMI-1640 complete medium (Sigma-Aldrich, Milan, Italy) added with GMC-SF (100 ng/mL) and IL-4 (200 ng/mL) (PeproTech, Rocky Hill, NJ, USA) for 5 days. On the 6^{th}

day, antigens were added (0.22 mg/mL), and the cells were cultured for48 hours. As positive control to activation, LPS (100 ng/mL) was used.

DCs phenotyping: For the phenotype analysis both active and control DCs were incubated with mAbs for 30 min and washed with PBS containing 0.1% bovine serum albumin (BSA) and 0.1% sodium azide. DCs were labeled with mAbs for CD80 FITC, CD83 PE, CD86 APC (BD Pharmingen, San Jose, CA, USA). Samples were analyzed by FACS Canto II cytometer (Becton-Dickinson, San Jose, CA, USA) using the DIVA software. Native AlaProAspThr(Tn)ArgPro was purchased from Giotto Biotech.

Cell viability: Viability of 4T1 cells and PBMC was quantified by PrestoBlue® assay (Thermo Fisher Scientific). Cells were seeded (5,000 cells/well, in triplicate) into 96-well plates. Cells were treated with serial concentrations of **mime[4]CRM**; **Glc-CRM** or CRM_{197} and incubated for up to 96 hours. After, part of the medium was removed, and PrestoBlue® reagent was added for up to 60 minutes at 37 °C. During this time, PrestoBlue® reagent is modified by the reducing environment of the viable cell and turns red in color, becoming highly fluorescent. Color change was detected using fluorescence Excitation 535/Emission 615 nm using Microplate Reader InfiniteF200 (Tecan). Experiments were repeated three independent times.

Cell culture: 4T1-luc, Mouse Mammary Gland Adenocarcinoma Cell Line (stably expressing firefly Luciferase gene) was cultured in RPMI-1640 medium (ECB9006L, Euroclone, Milan, Italy) supplemented with 10% fetal bovine serum (ECS0180D, Euroclone, Milan, Italy), 2 mM L-Glutamine (ECB3000D, Euroclone), and 1% antibiotics (ECB3001D, 10,000 U/mL

penicillin, 10 mg/mL streptomycin, Euroclone) at 37 °C in a 5% CO_2 incubator.

Tumor implantation and Bioluminescence imaging (BLI): 4T1-luc cells (8 x 104 cells in 50 µL PBS) were implanted into the lower right inguinal mammary fat pad (using insulin syringes with 32 G needles) of female (n=19) immunocompetent syngeneic mice model (4 weeks old; strain: BALB/c). Progression of tumor cells was monitored weekly by non-invasive *in vivo* bioluminescence imaging (BLI), using IVIS 3D Illumina Imaging System (Xenogen/Caliper) as previously described.[146] For *in vivo* BLI acquisitions, mice were anesthetized using 1-3% isoflurane (B104F16C, Piramal/Healthcare) and intraperitoneally injected (100 µL per 10 g body weight) with a 15 mg/mL solution of D-luciferin (122799, Perkin Elmer) in PBS. After 10 minutes from D-Luciferin injection, BLI acquisitions were performed using the auto-exposure mode. For tumor growth curve, the total flux of photons (photons per sec) within each region of interest (ROI) from bioluminescence image were quantified using the Living Images Software Package 3.2 (Xenogen-Perkin Elmer). Four-week-old female BALB/c mice were purchased from Envigo (Italy). Mice were housed in individually ventilated cages in an experimental room. The animals were maintained under standard conditions of light (12 hour-light/12 hour-dark), temperature (23 - 25 °C), humidity (50 to 60%) and rodent chow and water ad libitum. Survival analysis was performed using the Kaplan-Meier log-rank survival test with the SPSS16 statistical package. Ethical approval for mouse use released from Ministero della Sanità 546/2015 PR to the Director of Studies, M. Zollo, 19/6/2015, art. 31 D.Lgs. 26/2104.

***In vivo* drug administration:** For tumor growth evaluation and immune system evaluation, seven days after the implantation of tumor cells, mice (2 groups, n= 9 treated with **mime[4]CRM** and n=10 controls (CRM_{197}), grouped according to their BLI values were administered subcutaneously neck down the back with **mime[4]CRM** at a dose of 17 mg/kg/weekly resuspended in 100 µL PBS or with CRM_{197} as vehicle for the control group once weekly for 5 weeks. Mice were then sacrificed at days 28, 42 and at the end of the experiment (*i.e.* on day 54 from the tumor implantation). Serum, primary tumor and different organs were collected for the ELISA and immunohistochemistry analyses.

Immunofluorescence staining: Primary tumors were fixed in 4% paraformaldehyde (PFA) and embedded in paraffin. tissue sections (4-µm-thick) were deparaffinized in Xylene Substitute (A5597; Sigma Aldrich) 30 min and rehydrated in 100% - 90% -70%- 50% and 30% ethanol and washed with PBS 1X and PBS containing 0.02% TRITON-X 100 (215680010; Acros Organics). For antigen retrieval, the slides incubated in pre-warmed 10 mM sodium citrate buffer (pH 6.0) in microwave for 15 minutes, The slides were left to cool at room temperature for 30 min. to decrease the non-specific background fluorescence, the sections covered with 0.2% Trypsin (T2600000; Sigma-Aldrich) and 0.001% $CaCl_2$ and incubated for 10 minutes at 37 °C in humidified chamber, The tissue sections were washed with PBS and then blocked with a blocking solution containing 6% Bovine serum albumin (A9418; Sigma Aldrich) and 5% Fetal bovine serum (ECS0180L; Euroclone) 20 mM $MgCl_2$ in PBS containing 0.02% Triton-X 100 for 1 hour at room temperature with the following primary antibodies

overnight at 4 °C: anti-mouse CD4 FITC-conjugated (22150043, 1:200), FOXP3 (ab47285, 1:500), CD68 (ab53444,1:100), CD163 (sc-33560, 1:100), PD1/CD270 (E-AB-27294, 1:100), F4/80 (sc-52664, 1:100), CD11c/ Integrin alpha-x B-6 (cs-46676, 1:100). Tissue sections were washed in PBS 1X and PBS containing 0.02% TRITON-X 100 and were incubated with secondary antibody, anti-rabbit Alexa Fluor 546 (Thermo Fisher, #A10040, 1:200), anti-mouse Alexa Fluor 594 (ab150116, 1:200). DNA was stained with DAPI (#62254; Thermo Fisher). Confocal microscopy was carried out using a laser scanning confocal microscope (LSM 510 META, Zeiss), with the 63× oil immersion objective. Cell counting on immunofluorescence staining was performed using ImageJ software. The total number of cells were estimated by counting DAPI positive nuclei from three different images per sample. Later, the percentage of positive cells were determined as follow: (number of positive cells * 100)/ total number.

ELISA Tests: 96-well ELISA plates were coated with 0.5 µg/well of the hexapeptide AlaProAspDAPArgPro (as negative control) or with the glycosylated, AlaProAspDAP(**13**)ArgPro hexapeptide overnight at 4 °C. After washing with PBS containing 0,05% Tween 20 (PBST), wells were blocked with 200 µL/well of 5% BSA in PBS and incubated for 2 h at room temperature. Sera from immunized mice were added to wells as three-fold steps serial dilutions with PBST, compared to wells without sera. After washing with PBST, plates were incubated with HRP conjugated anti-mouse IgG (GE-Healthcare) at appropriated dilution for 2 hours at room temperature. For detection, plates were washed and added with 100 µL/well of 3,3',5,5'-tetramethylbenzidine (TMB - Sigma Aldrich) substrate

solution, and further incubated at room temperature for 30 minutes in the dark. Reactions were quenched by adding 100 µL/well of 2 M H_2SO_4 and immediately read at 450 nm using an Infinite F200 microplate reader (Tecan).

Flow Cytometry Analysis: 4T1 cells were grown to sub-confluency, detached from wells by enzyme-free Cell Dissociation Buffer (Gibco) and collected, then resuspended in PBST containing 5% BSA. Sera from immunized mice (1:100 dilution) were added to cells and incubated on ice for 30 minutes. After washing with PBST-BSA, cells were incubated with APC conjugated donkey anti-mouse IgG (Thermo Fisher Scientific) and incubated on ice for 30 minutes. After washing, stained cells were measured using BD FACSCanto™ II, and data analysis was performed using the FlowJo software.

Th1/Th2/Th17 cytokine profile assessment: Serum samples from mice treated with **mime[4]CRM** or from control mice were analyzed for the presence of Th1, Th2, and Th17 cytokines by using RayBio Quantibody Mouse TH17 Array 1 (RayBio, Norcross, GA), according to the manufacturer's instructions. Eighteen different cytokines were evaluated: interleukin 1 beta (IL-1β), IL-2, IL-4, IL-5, IL-6, IL-10, IL-12p70, IL-13, IL-17, IL-17F, IL-21, IL-22, IL-23, IL-28, interferon gamma (IFN-γ), macrophage inflammatory protein 3 alpha (MIP-3α), transforming growth factor beta 1 (TGFβ1), and tumor necrosis factor alpha (TNF-α). In brief, slides were incubated with blocking solution (RayBio) at room temperature for 30 min. Afterwards, standard cytokines and diluted sera (1:2) were incubated at room temperature for 2 hours. Slides were washed 5 times and probed

for 1 h at room temperature with Biotinylated Antibody Cocktail. After washing steps at room temperature, the slides were incubated 1 h at room temperature with Cy3 Equivalent Dye-Streptavidin, washed and finally dried by a compressed nitrogen. Signals were scanned on a ScanArray Plus scanner, using Imagene 8.0 software. Results were analyzed by Prism Graph Pad software. The statistical significance of the differences between groups was determined with Student's two-tailed t-test, while intra-group analysis was determined with two-way ANOVA test. $P < 0.05$ was considered to be statistically significant.

Human *ex vivo* tests: Patients: Twelve patients with breast cancer (BC) were enrolled for the study. All patients, that not received chemotherapy, underwent blood sampling and surgical resection of the primary lesion; patients with evidence of serious illness, immunodeficiency, or autoimmune or infectious diseases were excluded. Patients were enrolled after obtaining informed consent and approval of the local ethical committee.

Detection of mime[4]CRM-specific T cells in the peripheral blood of patients with BC: To assess the presence of **mime[4]CRM** specific T cells in the peripheral blood of BC patients, PBMCs were re-suspended in medium supplemented with 3% human serum. PBMCs (3x105) were cultured for 96 h in the presence of medium alone, **mime[4]CRM** (10 µg/mL), or **Glc-CRM** (as control, 10 µg/mL). At 16 h before harvesting, 0.5 µCi/well [3H]-thymidine (Amersham Pharmacia Biotech, Sweden) was added, and the radionuclide uptake was measured by a β-counter. The mitogenic index (MI) was calculated as the ratio between mean values of

cpm (counts for minute) obtained in **mime[4]CRM** or **Glc-CRM** stimulated cultures and those obtained in the presence of medium alone. A MI ≥3 was considered as a positive result.

Generation of T cell clones (Tcc) from TILs of the neoplastic tissue: Surgical specimens of BC tissue were dissociated in order to isolate TILs. First, tissue pieces were dissociated with the Tumor Dissociation Kit, human (Miltenyi Biotec) in combination with the gentleMACS™ Octo Dissociator (Miltenyi Biotec), to obtain a gentle and rapid generation of single-cell suspensions. Then, TILs were magnetically isolated with anti-human CD3 microbeads (Miltenyi Biotec) and cloned under limiting dilution. Briefly, single T-cell blasts were seeded in microwells (0.3 cells/well) in RPMI 1640 culture medium (SERO-Med GmbH) supplemented with 10% FCS HyClone (Gibco Laboratories), in the presence of 2 x 105 irradiated (9000 rad) PBMC, phytohemagglutinin (PHA) (0.5% vol/vol), and recombinant human interleukin-2 (50 U/mL) (PeproTech). At weekly intervals, 2 x 105 irradiated PBMC and IL-2 were added to each micro-culture to maintain the expansion of growing Tcc

Phenotypic characterization of isolated Tcc: Tcc' surface markers (CD3, CD4, CD8) expression was analyzed by flow cytometry using FACS Canto II cytometer (Becton-Dickinson) and the DIVA software. A total of 104 events for each sample was acquired.

Evaluation of mime[4]CRM-specific T cells in the neoplastic tissue of BC: To assess the presence of **mime[4]CRM**-specific T cells in the neoplastic tissue of BC, Tcc were re-suspended in medium supplemented with 3% human serum. Tcc were cultured for 48 h with irradiated autologous

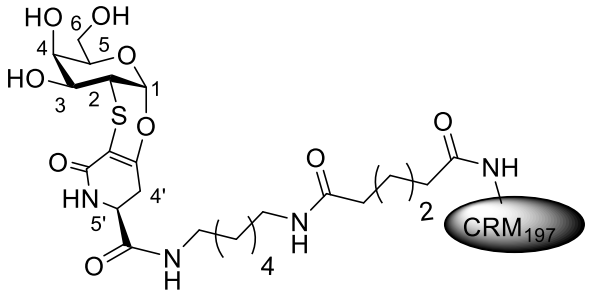
mononuclear cells in the presence of medium alone, **mime[4]CRM** (10 µg/mL or Glc-CRM (as control, 10 µg/mL). At 16 h before harvesting, 0.5 µCi/well [3H]-thymidine (Amersham Pharmacia Biotech) was added, and the radionuclide uptake was measured by a β-counter. The mitogenic index (MI) was calculated as the ratio between mean values of cpm (counts for minute) obtained in **mime[4]CRM** or Glc-CRM stimulated cultures and those obtained in the presence of medium alone. A MI ≥ 3 was considered as a positive result.

7.4.2 Synthesis

General method

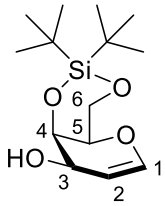
Analytical grade solvents and commercially available reagents were used without further purification. For anhydrous reactions, solvent stored over 3 Å molecular sieves were used. Silica gel flash column chromatography purifications were performed using Geduran® Si 60 (0.040-0.063 mm) or using the Biotage Isolera system and SNAP silica cartridges. TLC analyses were performed on glass Merck silica gel 60 F_{254} plates. ^{1}H NMR, ^{13}C NMR and 2D-NMR spectra were recorded on a 500 MHz Bruker AVANCE II at 298 K, unless otherwise stated. All chemical shifts are reported in parts per million (δ) referenced to residual nondeuterated solvent. Multiplicity abbreviation: b = broad, s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet were used. ESI-MS spectra were carried out on a linear ion-trap double quadrupole mass spectrometer using electrospray ionization (ESI) technique (LTQ-XL - Thermo Fisher) . Optical rotation measurements were performed on a JASCO DIP-370 polarimeter. Melting point were recorded on a STUART SMP3 version 2.0 or a BUCHI 510.

Synthesis of CRM_{197} glycoconjugates: general procedure



A 0.8 mM CRM_{197} solution in 150 mM NaPi, pH 7.2 buffer was treated with a 100-200 equivalents of activated carbohydrate previously dissolved in a minimal amount of 150 mM NaPi, pH 7.2 buffer. The reaction mixture was incubated at 4°C for 24 hours under shaking. The success of the reaction was proved by an SDS-PAGE. The conjugate was then purified from unreacted sugar by multistep washings with water using a 10 kDa MWCO membrane centrifugal device (Millipore). The number of attached moieties was estimated by UltraFlex III MALDI-TOF/TOF instrument (Bruker Daltonics) in linear mode and with positive ion detection.

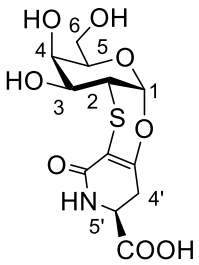
Synthesis of compound 11



To a solution of D-(+)-galactal (2.26 g, 15.47 mmol) in dry DMF (45 mL) cooled to-45 °C *[(tBu)₂Si(OTf)₂]* (5.5 mL, 16.8 mmol) was added dropwise. The solution was stirred at -45 °C for 1 h and then pyridine (pH ≃ 7) was added for quenching. The mixture was diluted with EtOAc and washed with $NaHCO_3$ s.s. (x3) and brine (x2). The organic layer was dried over Na_2SO_4, filtered, and the solvent was removed under vacuum to give crude **11** (4.75 g) as a yellowish solid that was used without further purification. The spectroscopic and analytical data were in agreement with those reported previously.[82]

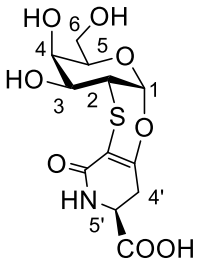
^{1}H NMR (500 MHz, $CDCl_3$) δ: 6.30 (dd, 1H, $J_{1,2}$ = 6.5 Hz, $J_{1,6}$= 1.4 Hz, H-1), 4.70 (m, 1H, H-2), 4.41-4.38 (m, 1H, H-4), 4.36-4.30 (m, 1H, H-3), 4.29-4.20 (m, 2H, H-6), 3.87 (m, 1H, H-5), 2.74 (d, 1H, J_{OH-3} = 11.4 Hz, OH), 1.07 (s, 9H, $(CH_3)_3$C-Si), 1.00 (s, 9H, $(CH_3)_3$C-Si).

Synthesis of compound 13 from 26



To a solution of **26** (100 mg, 0.24 mmol) in THF (1 mL), H_2O (100 µL) and Pd/C (10 wt.%, 10 mg) were added. The mixture was stirred at room temperature under H_2 atmosphere for 2 h and then filtered through a pad of Celite® and eluted with CH_3OH. The filtrate was further filtered through a 0.22 µm syringe filter. Evaporation of the solvent under vacuum gave a crude product that was purified by trituration with a small amount of CH_3OH and Et_2O. Pure **13** was isolated as a white solid (72 mg, > 90%).

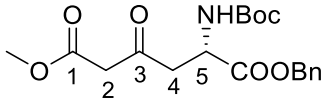
Synthesis of compound 13 from 27



A solution of **27** (100 mg, 0.22 mmol) in a solution of NH_3 (2M) in CH_3OH (1 mL) was stirred overnight at room temperature. After completion, evaporation of the solvent under vacuum gave a crude product that was purified by trituration with a small amount of CH_3OH and Et_2O. Pure **13** was isolated as a white solid (68 mg, > 90%). The spectroscopic and analytical data were in agreement with those reported previously.[82]

ESI-MS m/z (%): 332.09 (100) $[M]^-$, 665.09 (65) $[2M]^-$, **^{1}H NMR** (500 MHz, D_2O) δ: 5.83 (d, $J_{1,2}$ = 2.8 Hz, 1H, H-1), 4.24-4.19 (m, 1H; H-5), 4.17-4.12 (m, 1H; H-5'), 4.07-4.03 (m, 1H; H-4), 3.87-3.75 (m, 3H; H-3, H-6), 3.46 (dd, $J_{2,1}$ = 2.8 Hz, $J_{2,3}$ = 11.3 Hz, 1H, H-2), 2.81-2.75 (A part of an ABX system, $J_{A,X}$ = 6.0 Hz, $J_{A,B}$ 16.8 Hz, 1H; H-4'$_a$), 2.71-2.63 ppm (B part of an ABX system, $J_{B,X}$ = 8.7 Hz, $J_{B,A}$ = 16.8 Hz, 1H; H-4'$_b$).

Synthesis of compound 19

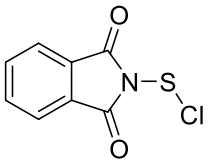


Under N_2 atmosphere, N-Boc-Asp-OBn (10.01 g, 30.96 mmol) was dissolved in dry CH_2Cl_2 (30 mL), DMAP (4.30 g, 35.19 mmol) and Meldrum acid (6.70 g, 46.52 mmol) were added, the solution was cooled to 0 °C and a solution of DCC (7.30 g, 35.38 mmol) in dry CH_2Cl_2 (20 mL) was added. The suspension was kept at 0 °C for 24 h. After completion, the precipitate was removed by filtration and washed with CH_2Cl_2 (x3). The filtrate was washed with $KHSO_4$ 1 M (x3) and brine (x1), dried over Na_2SO_4, filtered, and the solvent was removed under vacuum to give crude **25** (16.75 g) as a yellowish solid that was used without further purification. Meldrum's acid adduct **25** was solubilized in a mixture Toluene/CH_3OH 4:1 and the reaction mixture was heated at reflux for 17 h. Evaporation of the solvent under reduced pressure gave a crude product which was purified by flash column chromatography on silica gel (EtOAc/EP 4:1 -> 2:1) to give pure **19** as a white solid (10.21 g, 87% over 2 steps). The spectroscopic and analytical data were in agreement with those reported previously.[81]

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.40-7.31 (m, 5H, Bn), 5.48 (d, J = 8.3 Hz, 1H, NH), 5.17 (s, 2H, CH_2Ph), 4.61-4.55 (X part of an ABX system, $J_{X,A}$ = 10.5, $J_{X,}$

$_B$ = 4.2 Hz, $J_{X,NH}$ = 9.5 Hz, 1H, H-5), 3.73 (s, 3H, OCH_3), 3.45 (s, 2H, CH_2, H-2), 3.31-3.25 (A part of an ABX system, $J_{A,B}$ = 18.3 Hz, $J_{A,X}$ = 4.4 Hz, 1H, H-4_a), 3.16-3.10 (B part of an ABX system, $J_{B,A}$ = 18.3 Hz, $J_{B,X}$ = 4.4 Hz, 1H, H-4_b), 1.44 (s, 9H, $C(CH_3)_3$), **^{13}C NMR** (125 MHz, $CDCl_3$) δ: 201.2, 171.4, 167.5, 156.0, 135.9, 129.2, 129.0, 128.8, 80.8, 68.1, 53.1, 50.3, 49.6, 45.5, 29.0.

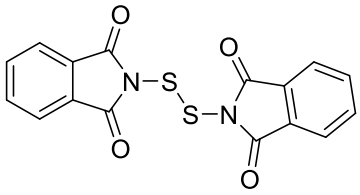
Synthesis of compound 20



To a solution of N,N'-dithiobis(phthalimide) **23** (12.45 g, 34.93 mmol) in 300 mL of HPLC-grade CH_2Cl_2, sulfuryl chloride (100 mL) was added via a dropping funnel at 0 °C. The yellow reaction mixture was stirred for several days (2-4) until completion monitored by NMR. All volatile materials were removed under vacuum, affording the title compound **20** with an NMR purity of 94% (13.83 g, 93%). The spectroscopic and analytical data were in agreement with those reported previously.[147]

^{1}H NMR ($CDCl_3$, 500 MHz) δ: 8.09-8.04 (m, 4H), 7.94-7.90 (m, 4H)

Synthesis of compound 23

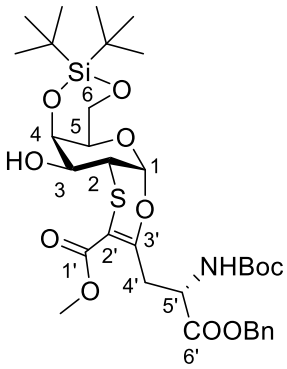


A solution of phthalimide (5.00 g, 33.98 mmol) and Et_3N (4.15 g, 41.01 mmol) in THF (100 mL) was cooled to -20 °C and S_2Cl_2 (2.31 g, 17.11 mmol) was added dropwise. After 2 h at room temperature, 200 mL of water were added, and the precipitate was recovered by filtration. The solid was washed with a minimum amount of Et_2O and recrystallized from a $CHCl_3/CH_3OH$ (2:1) to give white needles of N,N'-dithiobis(phthalimide) **23** (4.97 g, 82%). The

spectroscopic and analytical data were in agreement with those reported previously.[147]

^{1}H NMR ($DMSO_{d6}$, 500 MHz) δ: 8.05-8.01 (m, 4H), 8.00-7.96 (m, 4H)

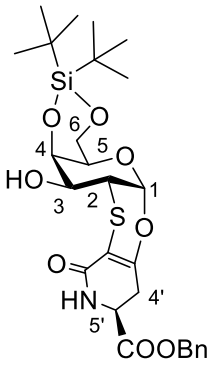
Synthesis of compound 24



To a solution of crude **11** (4.76 mmol) in Cl_2Et (10 mL), the β-ketoester **19** (2.43 g, 6.40 mmol), pyridine (4.5 mL) and PhtNSCl **20** (1.16 g, 5.44 mmol) were added, in the dark. The mixture was heated to 45 °C and stirred in the dark until completion. The precipitate was eliminated by filtration and the filtrate was diluted with CH_2Cl_2 and washed with a saturated solution of HCl 1M (x4). The organic layer was dried over Na_2SO_4 and the solvent was removed under vacuum. After impurity crystallization from Et_2O, the filtrate was concentrated to dryness and the residue was purified by flash chromatography on silica gel (EP/EtOAc 9:1 -> 8:2) to afford **24** (1.83 g, 55% over 2 steps) as a pale-yellow solid. The spectroscopic and analytical data were in agreement with those reported previously.[82]

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.38-7.31 (m, 5H, Bn), 5.63 (d, 1H, $J_{1,2}$ = 2.8 Hz, H-1), 5.39 (bd, 1H, NH), 5.22-5.07 (m, 2H, CH_2Ph), 4.70-4.57 (m, 1H, H-4'$_a$), 4.38 (m, 1H, H-4), 4.32-4.18 (m, 2H, H-6), 4.04 (m, 1H, H-5), 3.75 (s, 3H, OCH_3), 3.74-3.56 (m, 2H, H-3, H-5'), 3.24 (dd, 1H, $J_{2,1}$ = 2.8 Hz, $J_{2,3}$ = 10.6 Hz, H-2); 2.86 (1H, $J_{4'b,4'a}$ = 13.9 Hz, $J_{4'b,5'}$ = 4.3 Hz, HB-4'$_b$), 2.65 (bd, 1H, J_{OH-3} = 10.0 Hz, OH), 1.41 (s, 9H, $3CH_3$ - Boc), 1.05 (s, 9H, $(CH_3)_3$C-Si), 1.04 (s, 9H, $(CH_3)_3$C-Si).

Synthesis of compound 25

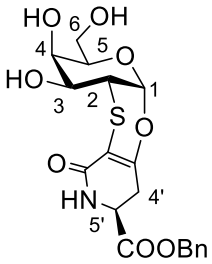


To a solution of **24** (2.19 g, 4.15 mmol) in dry CH_2Cl_2 (60 mL), TFA (3.2 mL, 41.58 mmol) was added. The solution was stirred at rt for 2 h, then the mixture was concentrated to dryness. The solid obtained was dissolved in $CHCl_3$ (60 mL) and Et_3N (7 mL) was added. The mixture was heated to 40 °C and stirred for 48 h. After completion, the mixture was diluted with $CHCl_3$ and washed with H_2O (x2), NH_4Cl (x2) and brine (x2). The organic layer was dried over Na_2SO_4, the solvent was removed under vacuum and the residue was purified by flash chromatography on silica gel (CH_2Cl_2/EtOAc 6:4) to afford pure **25** (1.49 g, 84% yield). The spectroscopic and analytical data were in agreement with those reported previously.[82]

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.39-7.31 (m, 5H, Bn), 6.04 (s, 1H, NH), 5.62 (d, $J_{1,2}$ = 2.6 Hz, 1H, H-1), 5.28-5.22 (A part of an AB system, $J_{A,B}$ = 12.0 Hz, 1H, CH_2Ph), 5.21-5.15 (B part of an AB system, $J_{B,A}$ = 12.0 Hz, 1H, CH_2Ph), 4.65-4.59 (X part of an ABX system, $J_{X,A}$ = 10.5, $J_{X,B}$ = 4.2 Hz, $J_{X,NH}$ = 9.5 Hz, 1H, H-5''), 4.50 (m, $J_{4,3}$ = 2.7 Hz, 1H, H-4), 4.30-4.27 (A part of an AB system, $J_{A,B}$ = 12.5 Hz, 1H, H-6_a), 4.20-4.15 (B part of an AB system, AB, $J_{B,A}$ = 12.8 Hz, 1H, H-6_b), 3.98 (bs, 1H, H-5), 3.69-3.95 (m, 1H, H-3), 3.47 (dd, $J_{2,1}$ = 2.6 Hz, $J_{2,3}$ = 11.0 Hz, 1H, H-2), 3.71-3.66 (A part of an ABX system, $J_{A,B}$ = 14.0 Hz, $J_{A,X}$ = 11.0 Hz, 1H, H-4''$_a$), 2.90-2.83 (B part of an AB system, $J_{B,A}$ = 14.2 Hz, $J_{B,X}$ = 4.3 Hz, 1H, H-4''$_b$), 1.67 (bs, 1H, OH),1.07 (s, 9H, $(CH_3)_3$C-Si), 1.05 (s, 9H, $(CH_3)_3$C-Si), **^{13}C NMR** (125 MHz, $CDCl_3$) δ:169.2 (CO), 164.7 (CO), 154.3 (C_q), 134.6 (C_q, Ph), 128.8 (CH, Ph), 128.8 (2CH, Ph), 128.5 (2CH, Ph),

97.9 (C_q), 96.4 (CH, C-1), 72.2 (C-4), 69.7 (C-5), 67.9 (CH_2Ph), 66.8 (C-6), 66.3 (C-3), 51.5 (C-5''), 39.7 (C-2), 30.9 (C-4''), 27.7 ($C(CH_3)_3$), 27.3 ($C(CH_3)_3$), 23.8 ($(CH_3)_3C$-Si), 20.8 ($(CH_3)_3C$-Si).

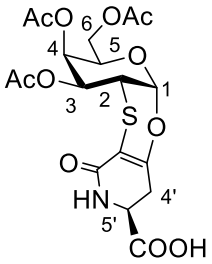
Synthesis of compound 26



To a suspension of **25** (830 mg, 1.47 mmol) in THF a 2M freshly solution of TBAHF (1.5 mL) was added and the reaction was stirred at room temperature. After 6 h volatile materials were removed under vacuum and the residue was purified by flash chromatography on silica gel (CH_2Cl_2/CH_3OH 9:1) to afford pure **26** as a white solid (600 mg, 97%). The spectroscopic and analytical data were in agreement with those reported previously.[82]

^{1}H NMR (500 MHz, CD_3OD) δ: 7.39-7.31 (m, 5H, Bn), 5.57 (d, $J_{1,2}$ = 2.8 Hz, 1H, H-1), 5.28-5.20 (m, 2H, CH_2Ph), 4.36-4.33 (m, 1H, H-5'), 4.04-4.00 (m, 1H, H-5), 3.97-3.94 (m, 1H, H-4), 3.82-3.70 (m, 2H, H-6), 3.61 (dd, $J_{3,4}$ =3.0 Hz, $J_{3,2}$ = 10.8 Hz, 1H, H-3), 3.45 (dd, $J_{2,1}$ =2.8 Hz, $J_{2,3}$ = 10.8 Hz, 1H, H-2), 3.08-3.02 (A part of an ABX system, $J_{4'a,5'}$ = 6.8 Hz, $J_{4'a,4'b}$ = 16.8 Hz, 1H, H-4'$_a$), 2.77-2.72 ppm (B part of an ABX system, $J_{4'b,5'}$ = 5.2 Hz, $J_{4'b,4'b}$ = 16.8 Hz, 1H, H-4'$_b$), **^{13}C NMR** (125 MHz, CD_3OD) δ: 170.7, 166.1, 156.1, 135.6, 128.1, 127.9, 127.7, 96.9, 95.9, 73.6, 68.9, 67.1, 65.9, 61.3, 51.1, 39.1, 30.3.

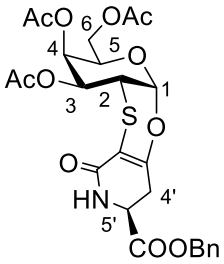
Synthesis of compound 27



To a solution of **28** (3.4 g, 6.18 mmol) in THF (70 mL), H_2O (1 mL) and Pd/C (10 wt.%, 320 mg) were added. The mixture was stirred at room temperature under H_2 atmosphere for 2 h and then filtered through a pad of Celite®. Evaporation of the solvent under vacuum gave 185 mg of crude product that was purified by flash chromatography on silica gel (CH_2Cl_2 /CH_3OH 9:1 -> 8/2) to give pure **27** as a white solid (2.8 g , > 98%).

$[\alpha]^{25}_D$= +74.3 (c 0.03 in CH_3OH), **^{1}H NMR**(500 MHz, CD_3OD) δ: 5.78 (d, $J_{1,2}$ = 2.8 Hz, 1H, H-1), 5.44 (m, 1H, H-4), 5.23 (dd, $J_{3,2}$ = 11.7 Hz, $J_{3,4}$ = 2.7 Hz, 1H, H-3), 4.60 (m, 1H, H-5), 4.22-4.17 (A part of an ABX system, 1H, H-6_a), 4.17-4.12 (B part of an ABX system, 1H, H-6_b), 4.11-4.06 (m, 1H, H-5'), 3.72 (dd, $J_{2,1}$ = 2.8 Hz, $J_{2,3}$ = 11.6 Hz, 1H, H-2), 2.91-2.85 (A part of an ABX system, $J_{A,X}$ = 6.6 Hz, $J_{A,B}$ 16.8 Hz, 1H; H-4'$_a$), 2.76-2.68 (B part of an ABX system, $J_{B,X}$ = 8.0 Hz, $J_{B,A}$ = 16.8 Hz, 1H; H-4'$_b$), 2.17 (s, 3H, CH_3), 2.05 (s, 3H, CH_3), 2.02 (s, 3H, CH_3), **^{13}C NMR** (125 MHz, CD_3OD) δ: 172.9 (CO), 170.4 (CO), 170.0 (CO), 169.8 (CO), 166.6 (CO), 155.5 (C_q), 97.0 (Cq), 95.9 (C-1), 69.1 (C-5), 67.1 (C-4), 65.5 (C-3), 61.5 (C-6), 51.3 (C-5'), 36.6 (C-2), 30.6 (C-4'), 20.7 (CH_3), 20.6 (CH_3), 20.6 (CH_3).

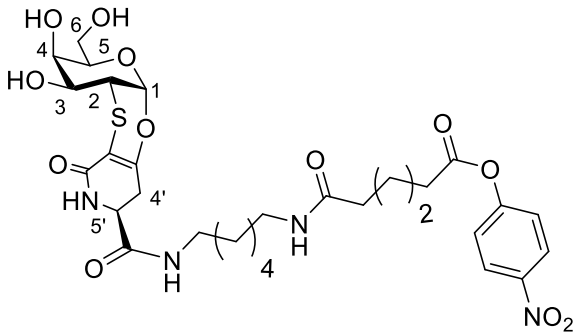
Synthesis of compound 28



Derivative **26** (500 mg, 1.19 mmol) was dissolved in CH_2Cl_2 (10 mL) and pyridine (4 mL, 49.5 mmol), Ac_2O (2 mL, 14.4 mmol) and DMAP (15 mg, 0.12 mmol) were added. The reaction mixture was stirred at room temperature for 10 min and then diluted with CH_2Cl_2 and washed with HCl 1M (x5) and NaHCO3 (x2). The organic layer was dried over Na_2SO_4, filtered, and concentrated to dryness. The crude was purified by flash chromatography on silica gel (EP/EtOAc 1:1 -> 3/7) to give **28** (545 mg, 91%) as a white solid.

$[\alpha]^{25}_D$ = +83.3 (c 0.75 in $CHCl_3$), **^{1}H NMR** (500 MHz, $CDCl_3$) δ: 7.45-7.35 (m, 5H, Bn), 5.95 (bs, 1H, NH), 5.69 (d, $J_{1,2}$ = 2.4 Hz, 1H, H-1), 5.45 (dd, $J_{4,3}$ = 3.0 Hz, $J_{4,5}$ = 1.4 Hz, 1H, H-4), 5.26 (m, 2H, CH_2Ph), 5.07 (dd, $J_{3,2}$ = 11.8 Hz, $J_{3,4}$ = 3.0 Hz, 1H, H-3), 4.44 (m, 1H, H-5), 4.43-4.37 (m, 1H, H-5'), 4.19-4.11 (m, 2H, H-6_a, H-6_b), 3.66 (dd, $J_{2,1}$ = 2.4 Hz, $J_{2,3}$ = 11.8 Hz, 1H, H-2), 2.94-2.78 (m, 2H, H-$4'_a$, H-$4'_b$), 2.16 (s, 3H, CH_3), 2.06 (s, 3H, CH_3), 2.02 (s, 3H, CH_3), **^{13}C NMR** (125 MHz, $CDCl_3$): δ170.2 (CO), 169.8 (CO), 169.7 (CO), 168.9 (CO), 164.1 (CO), 153.8 (C_q), 134.7 (C_q, Ph), 128.8 (CH, Ph), 128.7 (CH, Ph), 128.4 (CH, Ph), 98.2 (C_q), 95.9 (C-1), 69.1 (C-5), 67.9 (CH_2Ph), 67.2 (C-4), 65.6 (C-3), 61.5 (C-6), 51.5 (C-1'), 36.6 (C-2), 30.8 (C-2'), 20.5 (Ac), 20.4 (Ac).

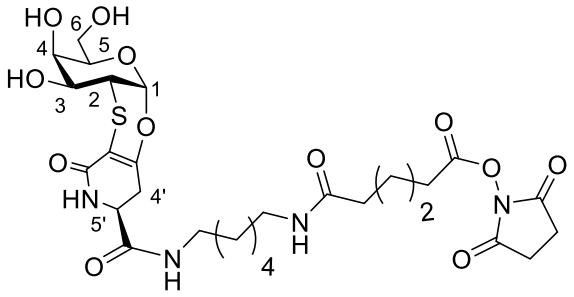
Synthesis of compound 29



To a solution of crude **34** (550 mg, 0.76 mmol) and NMM (300 µL, 2.73 mmol) in dry DMF (7.5 mL), bis(4-nitrophenyl)adipate (1.2 g, 1.76 mmol) in dry DMF (7.5 mL) was added and the mixture was stirred at room temperature overnight. After complete conversion, the solvent was evaporated and the residue purified by flash chromatography (CH_2Cl_2/CH_3OH 8:1) to give pure **29** (459 mg, 88% over two steps).

$[\alpha]^{22}_D$ = + 38.2 (c 0.385, CH_3OH), **ESI-MS m/z (%)**:703.25 (100) $[M+Na]^+$, 719.25 (50) $[M+K]^+$,**^{1}H NMR** (500 MHz, CD_3OD) δ: 8.32-8.27 (m, 2H), 7.40-7.36 (m, 2H), 5.67 (d, $J_{1,2}$ = 2.8 Hz, 1H, H-1), 4.14-4.10 (m, 1H, H-5'), 4.04-3.99 (m, 1H, H-5), 3.97-3.94 (m, 1H, H-4), 3.80-3.71 (m, 2H, H-6a, H-6b), 3.62 (dd, $J_{3,2}$ = 10.9 Hz, $J_{3,4}$ = 2.9 Hz, 1H, H-3), 3.46 (dd, $J_{2,1}$ = 2.8 Hz, $J_{2,3}$ = 10.9 Hz, 1H, H-2), 3.27-3.14 (m, 4H, $2CH_2$), 2.98-2.91 (m, 1H, H-4'a), 2.80-2.73 (m, 1H, H-4'b), 2.70-2.64 (m, 2H, CH_2), 2.28-2.23 (m, 2H, CH_2), 1.80-1.69 (m, 4H, $2CH_2$), 1.55-1.46 (m, 4H, $2CH_2$), 1.39-1.26 (m, 4H, $2CH_2$), **^{13}C NMR** (125 MHz, CD_3OD) δ: 175.6 (C_q), 172.7 (C_q), 172.6 (C_q), 167.6 (C_q), 157.8 (C_q), 157.2 (C_q), 146.8 (C_q), 126.1 (C_q), 123.9 (C_q), 98.3 (C-1), 96.8 (C_q), 75.0 (C-5), 70.2 (C-4), 67.2 (C-3), 62.6 (C-6), 53.2 (CC-5'), 40.4 (CH_2, C-2), 40.2 (CH_2), 36.7 (CH_2), 34.6 (CH_2), 32.1 (H-4'), 30.3 (CH_2), 30.2 (CH_2), 27.5 (CH_2). 27.3 (CH_2), 26.3 (CH_2), 25.3 (CH_2).

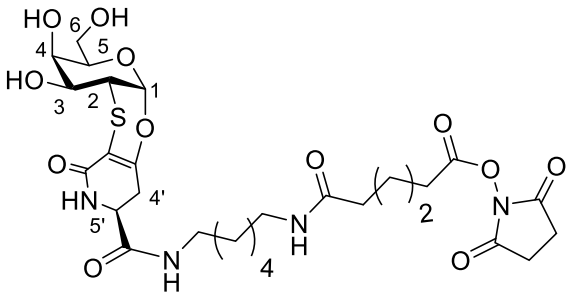
Synthesis of compound 30 from 29



Compound **29** (225 mg, 0.33 mmol) was dissolved in dry DMF (4 mL), solid NHS (114 mg, 0.99 mmol) was added and the mixture was stirred at room temperature overnight. After complete conversion, the solvent was evaporated and the crude was purified by precipitation from CH_3CN to give **30** as white solid (82 mg, 38%).

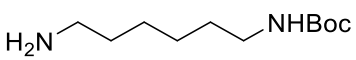
ESI-MS m/z (%): 679.75 (100) $[M+Na]^+$, **^{1}H NMR** (500 MHz, CD_3OD) δ: 5.67 (d, $J_{1,2}$ = 2.8 Hz, 1H, H-1), 4.18-4.14 (m, 1H, H-5'), 4.05-4.00 (m, 1H, H-5), 3.98-3.95 (m, 1H, H-4), 3.81-3.71 (m, 2H, H-6a, H-6b), 3.62 (dd, $J_{3,2}$ = 10.9 Hz, $J_{3,4}$ = 2.9 Hz, 1H, H-3), 3.46 (dd, $J_{2,1}$ = 2.8 Hz, $J_{2,3}$ = 10.9 Hz, 1H, H-2), 3.28-3.13 (m, 4H, $2CH_2$), 2.98-2.91 (m, 1H, H-4'a), 2.83 (s, 4H, CH_2), 2.81-2.75 (m, 1H, H-4'b), 2.69-2.63 (m, 2H, CH_2), 2.26-2.20 (m, 2H, CH_2), 1.77-1.68 (m, 4H, $2CH_2$), 1.55-1.45 (m, 4H, $2CH_2$), 1.39-1.30 (m, 4H, $2CH_2$), **^{13}C NMR** (125 MHz, CD_3OD) δ: 175.6 (C_q), 172.7 (C_q), 171.9 (C_q), 170.1 (C_q), 167.6 (C_q), 157.8 (C_q), 98.3 (C-1), 96.8 (C_q), 75.0 (C-5), 70.2 (C-4), 67.2 (C-3), 62.6 (C-6), 53.2 (C-5'), 40.5 (CH_2), 40.4 (C-2), 40.2 (CH_2), 36.5 (CH_2), 32.2 (H-4'), 31.3 (CH_2), 30.3 (CH_2), 30.2 (CH_2), 27.5 (CH_2). 27.3 (CH_2), 26.5 ($2CH_2$), 26.2 (CH_2), 25.2 (CH_2).

Synthesis of compound 30 from 34



To a solution of crude **34** (0.18 mmol) and NMM (170 µL, 1.55 mmol) in dry DMF (5 mL), bis(N-succinimidyl) adipate (221 mg, 0.65 mmol) was added and the mixture was stirred at room temperature overnight. After complete conversion, the precipitate was eliminated, and the solvent was evaporated. The residue was suspended in EtOAc and the precipitate was purified by flash chromatography (CH_2Cl_2/CH_3OH 8:1) to give pure **30** (60 mg, 51% over two steps)

Synthesis of compound 31

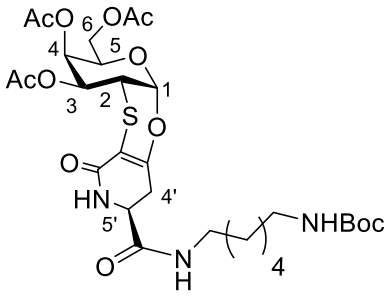


To a solution of 1,6-diaminohexane (40.0 g, 344 mmol) in CH_2Cl_2 (350 mL), cooled to 0 °C, a solution of di di-tert-butyl dicarbonate (15.0 g, 69 mmol) in CH_2Cl_2 (100 mL) was slowly added. After 16 h at room temperature, the mixture was diluted with CH_2Cl_2 and washed with Na_2CO_3 s.s. (x3) and H_2O (x1). The organic layer was dried over Na_2SO_4 and the solvent was removed under vacuum to give pure **31** as a colorless oil (14.64 g > 98%). The spectroscopic and analytical data were in agreement with those reported previously.[148]

ESI-MS m/z (%): 217.17 (100) $[M+H]^+$, **^{1}H NMR** (500 MHz, $CDCl_3$) δ: 5.15 (s, 2H, NH_2), 4.95 (bs, 1H, NH), 2.91 (dt, J = 6.5 Hz, J = 6.3 Hz, 2H, CH_2), 2.49 (t, J = 6.9 Hz, 2H, CH_2), 1.32-1.28 (m, 4H, 2 CH_2), 1.25 (s, 9H, 3 CH_3), 1.17-

1.11 (m, 6H, 3 CH_2),**^{13}C NMR** (125 MHz, $CDCl_3$) δ: 156.03 ; 78.96; 41.53; 40.43; 32.46; 29.95; 28.41; 26.49; 26.40.

Synthesis of compound 32

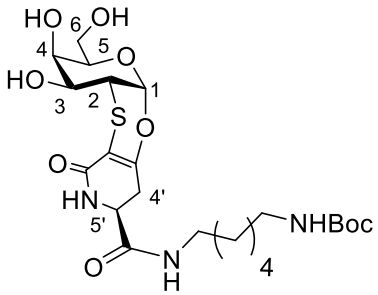


To a solution of **27** (425 mg, 0.93 mmol) in dry DMF (6 mL), a fresh solution of TBTU (597 mg, 1.86 mmol) and NMM (204 µL, 1.86 mmol) in dry DMF (12 mL) was added. Subsequently, N-Boc-1,6-hexanediamine (242 mg, 1.12 mmol) dissolved in dry DMF (2 mL) was added and the solution was stirred at room temperature for 3 h. After complete conversion, the solvent was evaporated and the residue purified by flash chromatography (CH_2Cl_2/CH_3OH 20:1) to give pure **32** (526 mg, 86%).

[α]$^{19}_D$ = +77.7 (c 0.60, $CHCl_3$), **ESI-MS *m/z* (%)**: 680.17 (100) $[M+Na]^+$, 696.17 (45) $[M+K]^+$, **^{1}H NMR** (500 MHz, $CDCl_3$) δ: 6.96 (bs, 1H, NH), 6.84 (bs, 1H, NH), 5.70 (d, $J_{1,2}$ = 2.7 Hz, 1H, H-1), 5.43-5.41 (dd, $J_{4,3}$ = 3.1 Hz, $J_{4,5}$ = 1.1 Hz, 1H, H-4), 5.02 (dd, $J_{3,2}$ = 11.7 Hz, $J_{3,4}$ = 3.1 Hz, 1H, H-3), 4.63 (bs, 1H, NH), 4.45-4.44 (m, 1H, H-5), 4.18-4.12 (m, 3H, H-5', H-6a, H-6b), 3.63 (dd, $J_{2,1}$ = 2.7 Hz, $J_{2,3}$ = 11.7 Hz, 1H, H-2), 3.30-3.20 (m, 2H, CH_2), 3.14-3.03 (m, 2H, CH_2), 3.02-2.84 (m, H-4'a, H-4'b), 2.15 (s, 3H, $COCH_3$), 2.05 (s, 3H, $COCH_3$), 2.01 (s, 3H, $COCH_3$),1.53-1.41 (m, 13H, $2CH_2$, tBu), 1.36-1.30 (m, 4H, $2CH_2$), **^{13}C NMR** (125 MHz, $CDCl_3$) δ: 170.5 (C_q) 170.1 (C_q), 170.0 (C_q), 169.5 (C_q), 165.2 (C_q), 156.3 (C_q), 155.6 (C_q), 96.8 (C_q), 95.9 (CH, C-1), 79.2 (C_q, tBu), 69.0 (C-5), 67.2 (C-4), 65.9 (C-3), 61.6 (C-6), 52.3 (C-5'), 40.1 (CH_2),

39.5 (CH_2), 36.3 (C-2), 30.8 (C-4'), 29.9 (CH_2), 29.1 (CH_2), 28.5 ($3CH_3$, tBu), 26.0 (CH_2), 25.8 (CH_2) 20.8 ($3CH_3$, CH_3).

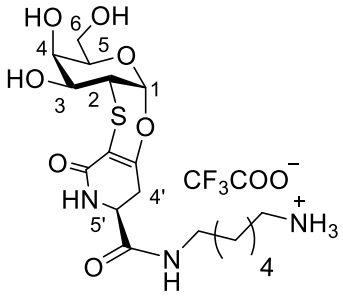
Synthesis of compound 33



Compound **32** (470 mg, 0.76 mmol) was dissolved in CH_3OH (8 mL), and NH_3 in CH_3OH 2M (4.5 mL) was added. After 2 h, the reaction mixture was concentrated under vacuum to give pure **33** (410 mg, >95%).

$[\alpha]^{22}_D$ = + 31.2 (c 0.345, CH_3OH), **ESI-MS m/z (%)**: 554.42 (70) $[M+Na]^+$, 570.42 (100) $[M+K]^+$, **^{1}H NMR** (500 MHz, CD_3OD) δ: 5.67 (d, $J_{1,2}$ = 2.8 Hz, 1H, H-1), 4.14-4.10 (m, 1H, H-5'), 4.03-3.99 (m, 1H, H-5), 3.96-3.94 (m, 1H, H-4), 3.80-3.70 (m, 2H, H-6a, H-6b), 3.61 (dd, $J_{3,2}$ = 11.0 Hz, $J_{3,4}$ = 2.9 Hz, 1H, H-3), 3.46 (dd, $J_{2,1}$ = 2.8 Hz, $J_{3,2}$ = 11.0 Hz, 1H, H-2), 3.28-3.16 (m, 2H, CH_2), 3.04 (m, 2H, CH_2), 2.93 (dd, $J_{4'a,4'b}$ = 16.7 Hz, $J_{4'a,5'}$ = 6.8 Hz, 1H, H-4'a), 2.78 (dd, $J_{4'b,4'a}$ 16.7 Hz, $J_{4b',5'}$ = 6.5 Hz, 1H, H-4'b), 1.43 (s, 13H, $2CH_2$, tBu), 1.36-1.30 (m, 4H, $2CH_2$), **^{13}C NMR** (125 MHz, CD_3OD) δ: 172.7 (C_q), 167.6 (C_q), 158.6 (C_q), 157.8 (C_q), 98.3 (C-1), 96.8 (C_q), 79.8 (C_q), 75.0 (CH, C-5), 70.2 (C-4), 67.2 (C-3), 62.6 (C-6), 53.2 (C-5'), 41.2 (CH_2), 40.5 (CH_2), 40.4 (C-2), 32.1 (H-4'), 30.9 (CH_2), 30.2 (CH_2), 28.8 ($3CH_3$, tBu), 27.4 ($2CH_2$).

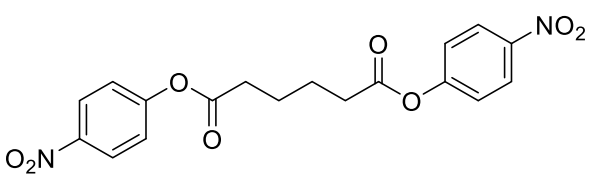
Synthesis of compound 34



To a suspension of **33** (410 mg, 0.77 mmol) in dry CH_2Cl_2 (13 mL), TFA (2 mL, 26.12 mmol) was added. After 2 h, the reaction mixture was concentrated under vacuum to give crude **34** (550 mg) that was used without further purification.

^{1}H NMR (500 MHz, CD_3OD) δ: 5.66 (d, $J_{1,2}$ = 2.8 Hz, 1H, H-1), 4.14-4.09 (m, 1H, H-5'), 4.04-3.99 (m, 1H, H-5), 3.95-3.93 (m, 1H, H-4), 3.80-3.71 (m, 2H, H-6a, H-6b), 3.61 (dd, $J_{3,2}$ = 10.9 Hz, $J_{3,4}$ = 2.9 Hz, 1H, H-3), 3.46 (dd, $J_{2,1}$ = 2.8 Hz, $J_{2,3}$ = 10.9 Hz, 1H, H-2), 3.29-3.17 (m, 2H, CH_2), 3.02-2.96 (m, 1H, H-4'a), 2.94-2.89 (m, 2H, CH_2), 2.80-2.74 (m, 1H, H-4'b), 1.70-1.62 (m, 2H, CH_2), 1.58-1.50 (m, 2H, CH_2) 1.47-1.32 (m, 4H, $2CH_2$).

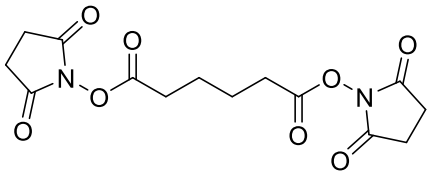
Synthesis of compound 35



To a suspension of 4-Nitrophenol (15.3 g, 110 mmol) in dry CH_2Cl_2 (150 mL), cooled to 0 °C, N,N-Diisopropylethylamine (14.2 g, 110 mmol) and adipoyl chloride (10.0 g, 54.63 mmol) were added. After 3 h at room temperature the mixture was diluted with CH_2Cl_2 and washed with H_2O (x3). The organic layer was dried over Na_2SO_4 and the solvent was removed under vacuum to give a crude that was purified by several washing with cold EtOAc to give pure **35** as a white solid (16.5 g, 78%).

^{1}H NMR: (500 MHz, DMSO-d_6) δ 8.31 (m, 4H, 2 CH), 7.46 (m, 4H, 2 CH), 2.72-2.70 (m, 4H, 2 CH_2), 1.78-1.76 (m, 4H, 2 CH_2).

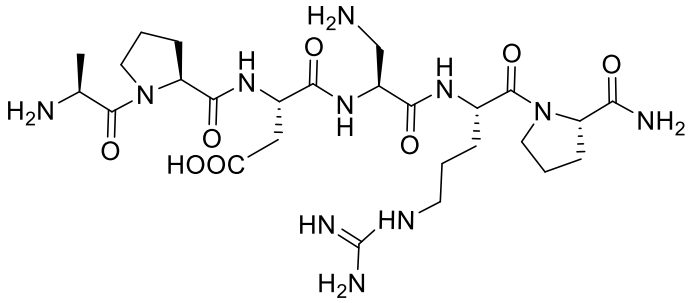
Synthesis of compound 36



To a solution of N-hydroxysuccinimide (12.7 g, 110 mmol) in THF (150 mL), cooled to 0 °C, N,N-diisopropylethylamine (14.2 g, 110 mmol) and adipoyl chloride (10.0 g, 54,63 mmol) were added. After 1 h at room temperature the suspension was filtered, and the precipitate was washed with hot iPrOH and THF. The product was purified by several crystallization from acetone to give white needles of **36** (13.0 g, 70%). The spectroscopic and analytical data were in agreement with those reported previously.[149]

^{1}H NMR: (500 MHz, $DMSO_{d6}$) δ: 2.82 (s, 8H), 2.76-2.72 (m, 4H).

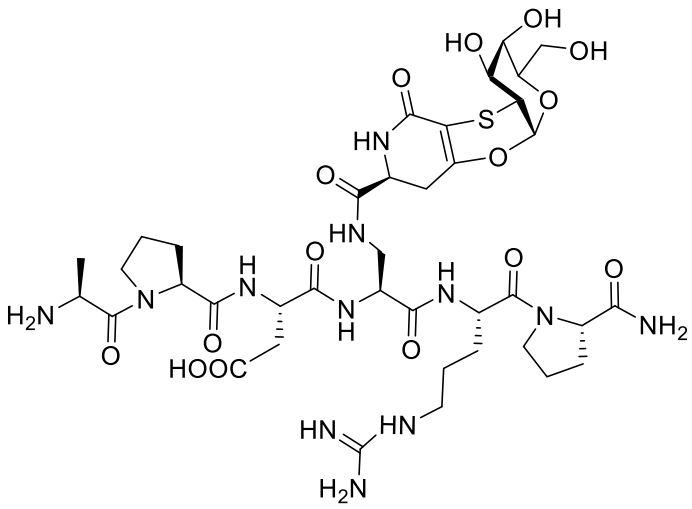
Synthesis of hexapeptide 37



Compound **37** was synthesized by following standard Fmoc solid-phase method. Starting from Fmoc-Pro-Rink-Amide MBHA resin. Sidechain deprotection and cleavage from the resin were achieved in a standard single step acidolysis reaction. Hexapeptide **37** was characterized by NMR and MS.

[α]^{D}22 = -71.4 (c 0.22, H2O); **ESI-MS m/z (%):** 320.81 (100) $[M+2H]^{2+}$, 640.38 (70) $[M+H]^{+}$, 662.38 (20) $[M+Na]^{+}$

Synthesis of glycosyl hexapeptide 38



Glycopeptide **38** was synthesized by coupling of derivative **27** to the protected and linked to the resin peptide, with PyBOP and DIPEA in DMF. After 2 h, the resin was washed with DMF and the acetylated glycopeptide was detached and deprotected by acidolysis reaction (TFA/TIS/H_2O 95:2.5:2.5) followed by treatment with NH_3 in CH_3OH 1 M. Glycopeptide **38** was purified by C18 reverse-phase chromatography (H_2O/CH_3CN 80:20) and characterized by NMR and HRMS.

ESI-HRMS m/z (%): calculated $[M+H]^{+}$ = 955,3938; found = 955.39298 (100) $[M+H]^{+}$

CHAPTER

8

Silica nanoparticles decorated with a TnThr mimetic as tool for solid state NMR (SSNMR) investigation

8.1 Introduction

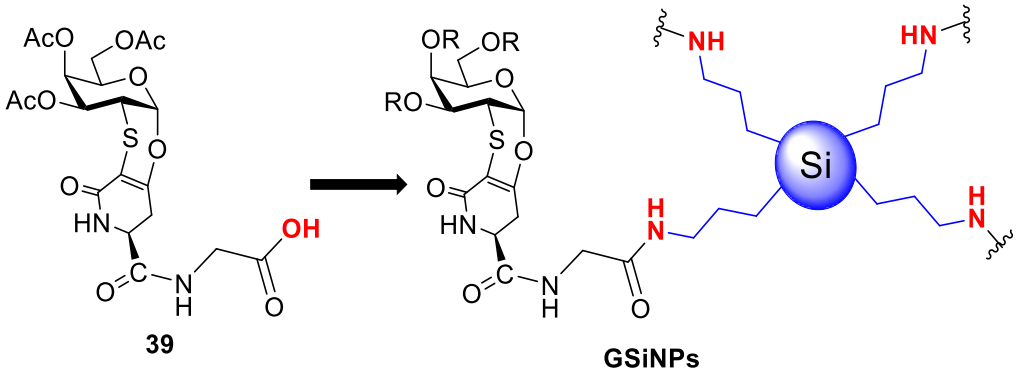
As widely discussed (**Chapter 5 - Section 5.2**), a successful workaround to face the low immunogenicity of TACAs is represented by multiple display, artificially obtained by their grafting onto multivalent scaffolds.[96,150] A plethora of natural, semi-synthetic and synthetic scaffolds was proposed for the multivalent presentation of selected glycosidic epitopes to the immune system like: peptides, proteins, dendrimers, oligonucleotides, and nanoparticles of various nature.[97,150,151] As already mentioned, glycosylated nanoparticles were successfully prepared using TnThr mimetic **13**. In particular, glycosyl superparamagnetic iron oxide nanoparticles (GMNPs)[98] and dextran-based single-chain polymer nanoparticles (DXT-SCPNs)[100] were synthesized and tested *in vitro*.

Macrophages and PBMCs stimulation observed for the multivalent constructs, with respect to the monovalent unit, confirmed the role of these nanoplatforms in mediating immune recognition. However, the characterization of such nanosystems proved to be demanding and sometimes rather challenging. Different advantages and limitations of each characterization techniques[152] complicate the choice and therefore a combinatorial characterization approach is needed. In this context, although NMR spectroscopy is a powerful analytical technique for the structural determination of soluble nanoscale materials, it suffers from two main problems: a) the sample shall be solubilized (task not always easy handling nanosystems) and b) a fast transverse relaxation rate that broaden the signals beyond the detection threshold. Solid state NMR (SSNMR), a cutting-edge technique, for the characterization of hybrid inorganic-biomolecular composites[153–159] can be used to overcome these obstacles. Surprisingly, at present no example of glycosyl-nanomaterial characterized by SSNMR is reported. During my PhD work, the versatility of SSNMR to investigate the organic coating surrounding nanoparticles was evaluated using silica nanoparticles (SiNPs) loaded with residues of TnThr mimetic **13**.[160] Easy to prepare and non-expensive, silica nanoparticles have interesting biomedical applications (*e.g.* site-specific delivery of drugs, imaging purposes) and received a great attention from the research community.[161–166]

8.2 Results and Discussion

8.2.1 Synthesis of TnThr-Silica nanoparticles (TnThr-SiNPs)

In this study[160], commercially available SiNPs presenting an organic coating were selected. In particular, SiNPs coated with aminopropyl chains (1 mmol/g NH_2 loading - 40-60 nm Ø), were chosen, because the suspected and matter of concern[167] toxic effects of SiNPs, is dramatically attenuated by surface-functionalization with amino groups.[168] In addition, the chain terminal amino group can be conveniently used to form an amidic linkage with the TnThr mimetic which is suitably functionalized with a spacer presenting a carboxylic group (**39** - **Scheme 8.2.1.1**).

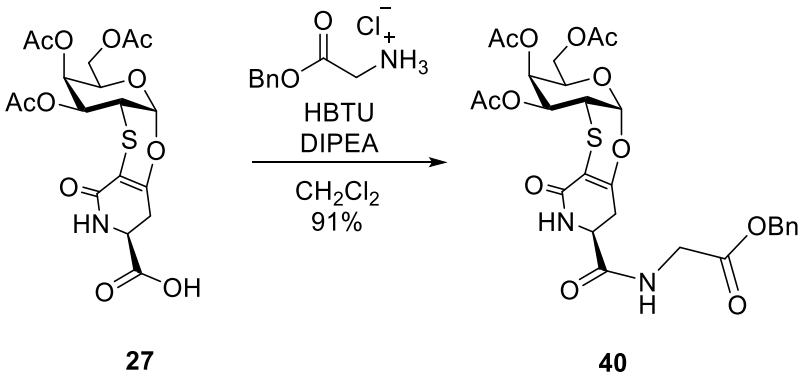


Scheme 8.2.1.1 Representative scheme of glycosylated silica nanoparticles (GSiNPs) preparation

For the preparation of the glycosylated silica nanoparticles (GSiNPs) decorated with **13**, glycine benzyl ester was used as spacer. The reasons for this choice are: *a)* the presence of easily derivatized amino- and carboxylic groups *b)* the biocompatibility of the glycosylated nanoparticles [although beyond the scope of this work, GSiNPs could find future

applications] and c) a convenient length to well present the mimetic graphed on the silica surface. Concerning this latter point, it is well known that too short spacers can cause an incorrect exposition of epitopes for adhesion phenomena onto nanoparticle surface; on the contrary, too long and non-polar alkyl chains have the inevitable tendency to aggregate. In both cases the presentation is disfavored. The glycine reasonably has the proper length.

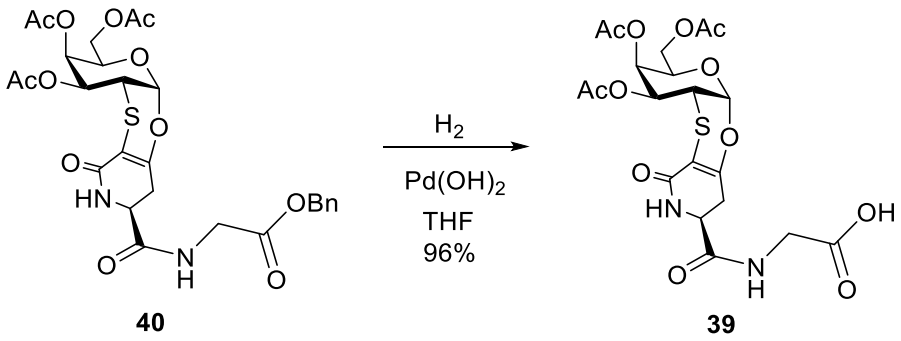
In the planned synthetic strategy, the first step is therefore the synthesis of compound **40**. Mimetic **27**, activated with O-(benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HBTU) in the presence of DIPEA, was reacted with glycine benzyl ester to afford, after flash chromatography, pure **40** in a 91% yield. (**Scheme 8.2.1.2**)



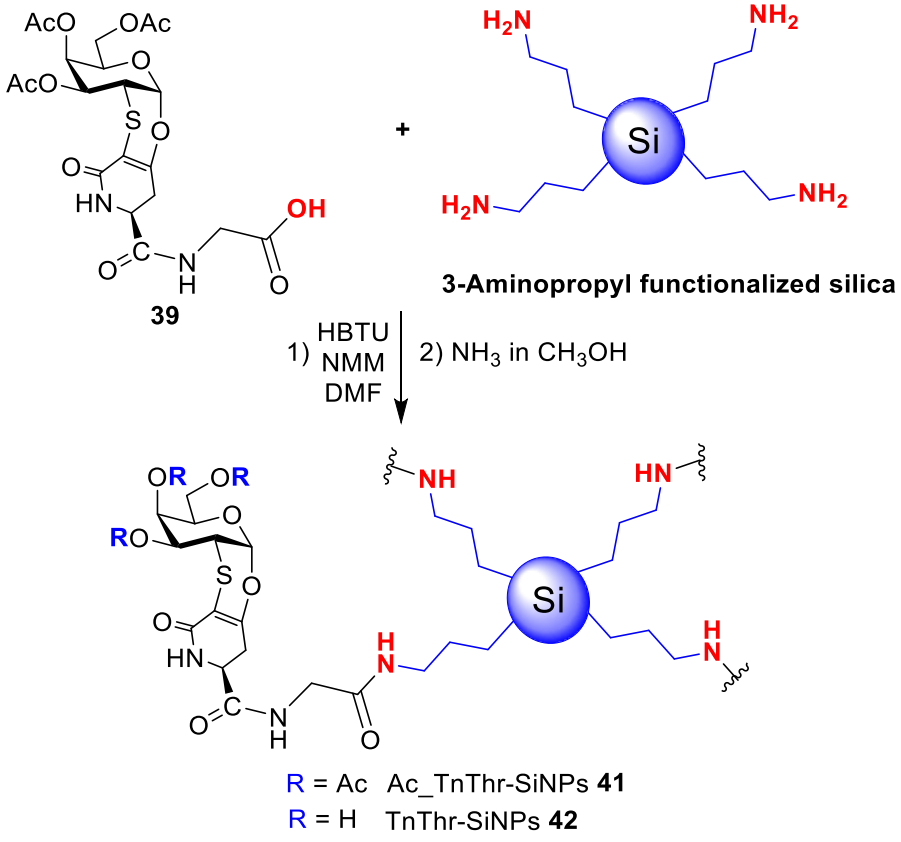
27 **40**

Scheme 8.2.1.2 Synthesis of amide **40**

Hence, **40** was converted in the derivative **39**, after a palladium catalyzed hydrogenation to remove the benzyl group (96% yield). (**Scheme 8.2.1.3**)



Scheme 8.2.1.3 Synthesis of carboxylic acid **39**



Scheme 8.2.1.4 Synthesis of glycosylated SiNPs (**GSiNPs**)

The amido-derivative **39** presenting an activated carboxylic group was then covalently linked to the propylamine-coated SiNPs by HBTU

mediated coupling. The acetylated **GSiNPs (Ac_TnThr-SiNPs** - **41**) were washed several times with DMF and CH_2Cl_2 to remove unreacted or adsorbed reagents. The subsequent deacetylation was performed on solid phase with a solution of NH_3 in CH_3OH (4M) and the nanoparticles obtained were washed with methanol to afford the desired glycosylated SiNPs (**TnThr-SiNPs** - **42**), (**Figure 8.2.1.4**).

8.2.2 SSNMR and ICP-MS characterization of GSiNPs

The loading of **GSiNPs** was qualitatively investigated by SSNMR. *{^{1}H}-^{13}C cross polarization (CP) under magic angle spinning (MAS) spectra* were recorded in isotopic natural abundance on *a)* propylamine-coated SiNPs (**Figure 8.2.2.1 - A**), *b)* acetylated glycosylated GSiNPs - **41** (**Figure 8.2.2.1 - B**), and *c)* deacetylated GSiNPs - **42** (**Figure 8.2.2.1 - C**). Signals of carbonyl (~170 ppm), anomeric (~100 ppm), and ethylene (~170 and ~100 ppm) carbons were observed only on the spectra of the glycosylated silica nanoparticles (**Figure 8.2.2.1 - B, C**), proving the successful grafting of the TnThr mimetic onto SiNPs. Noteworthy, the decrease in signal intensity observed in the spectrum of the deacetylated GSiNPs **42** for the signals of carbonyl and methyl carbon atoms, but not for those of the anomeric and ethylene carbons, proves the successful deacetylation of the TnThr mimetic.

The quantitative study of **GSiNPs** was carried out through Inductively Coupled Plasma (ICP) analysis. The amount of sulfur was dosed on an aliquot of nanoparticles and 36% of amino groups resulted correctly functionalized with TnThr mimetic **13**. The ICP data indicates a good

degree of functionalization, which could be probably further improve by increasing the equivalents of SiNPs used during the coupling reaction.

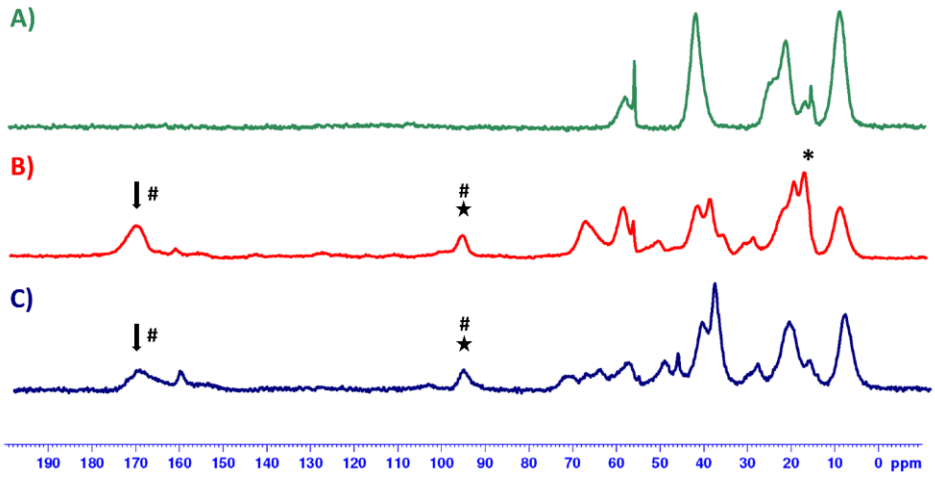


Figure 8.2.2.1 {^{1}H}-^{13}C CP MAS spectra of (**A**) propylamine-coated SiNPs, (**B**) acetylated glycosylated SiNPs **41** and (**C**) deacetylated glycosylated SiNPs **42** in 3.2 mm rotors, collected on 850 and 800 MHz spectrometers at 280 K and MAS of 14 kHz. The carbonyl, anomeric, ethylene and methyl signals are marked with arrows, stars, hashtags and asterisks, respectively.

8.3 Conclusions

In conclusion, an efficient glycosylation of SiNPs with the TnThr mimetic **13** has been reported. Glycosylated SiNPs were screened by {^{1}H}-^{13}C CP MAS spectra and ICP-MS analysis to prove, qualitatively and quantitatively, the successful coating of the nanoparticles. The results obtained with this model glyco-nanomaterial provide new perspective for the use of SSNMR for the characterization of biomolecules-coated nanoparticles, proving that it is possible to characterize the organic portion of hybrid organic and inorganic materials on solid state by using

SSNMR. Moreover, with this technique it is also possible to monitor the progress of a reaction performed on functionalized nanomaterials. Undeniably, it is important to combine different techniques to get as much information as possible for a complete structural elucidation; SSNMR can be one of these techniques.

8.4 Experimental section

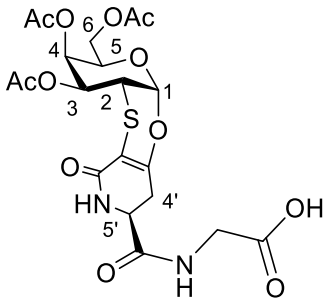
8.4.1 General methods

Analytical grade solvents and commercially available reagents were used without further purification. For anhydrous reactions, solvent stored over 3 Å molecular sieves were used. Silica gel flash column chromatography purifications were performed using Geduran® Si 60 (0.040-0.063 mm) or using the Biotage Isolera system and SNAP silica cartridges. TLC analyses were performed on glass Merck silica gel 60 F_{254} plates. ^{1}H NMR, ^{13}C NMR and 2D-NMR spectra were recorded on a 500 MHz Bruker AVANCE II at 298 K, unless otherwise stated. SSNMR experiments were performed on propylamine-coated SiNPs (Aldrich), acetylated glycosylated SiNPs and deacetylated glycosylated SiNPs. The samples were packed in 3.2 mm zirconia rotors. The $\{^{1}H\}$-^{13}C CP MAS spectra were recorded using the standard pulse sequence and parameters reported in literature. The spectra were collected at ~280 K on a Bruker AVANCE III 850 MHz wide-bore spectrometer (213.6 MHz ^{13}C Larmor frequency), equipped with 3.2 mm DVT MAS probe head in triple-resonance mode and on a Bruker AVANCE III 800 MHz narrow-bore spectrometer (201.2 MHz ^{13}C Larmor frequency) equipped with Bruker 3.2 mm Efree NCH probe-head, under MAS conditions of 14 kHz. All chemical shifts are reported in parts per million (δ) referenced to residual nondeuterated solvent. Multiplicity abbreviation: b = broad, s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet were used. ESI-MS spectra were carried out on a linear ion-trap double quadrupole mass spectrometer using electrospray ionization

(ESI) technique (LTQ-XL - Thermo Fisher) . Optical rotation measurements were performed on a JASCO DIP-370 polarimeter. Melting point were recorded on a STUART SMP3 version 2.0 or a BUCHI 510. ICP-MS analysis were performed by *Laboratorio di Microanalisi* (University of Florence).

8.4.2 Synthesis

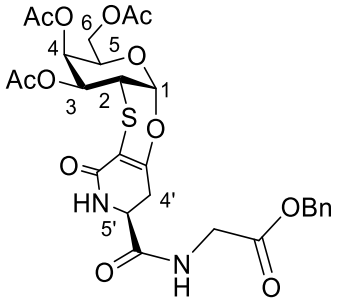
Synthesis of compound 39



To a solution of **40** (160 mg, 0.26 mmol) in THF/H_2O 100:1 (3 mL), $Pd(OH)_2/C$ (20 wt.%, 30 mg) was added. The reaction was stirred at room temperature for 16 h under H_2 atmosphere, diluted with EtOAc and filtered through a pad of Celite®. The filtrate was concentrated to dryness to give **39** (130 mg, > 96%) as a white solid.

$[\alpha]^{26}_D$ = +96.9 (c 0.0035, CH_3OD); **ESI-MS *m/z* (%):** 515.50 (100) $[M-H]^-$, 1031.33 (50) $[2M-H]^-$; **^{1}H NMR** (500 MHz, CD_3OD) δ: 5.79 (d, $J_{1,2}$ = 2.8 Hz, 1H, H-1), 5.42-5.39 (m, 1H, H-4), 5.10 (dd, $J_{3,2}$ = 11.7 Hz, $J_{3,4}$ = 2.8 Hz, 1H, H-3), 4.55-4.50 (m, 1H, H-5), 4.28-4.23 (m, 1H, H-5'), 4.20-4.10 (m, 2H, H-6a, H-6b), 4.00-3.88 (m, 2H, CH_2, Gly), 3.71 (dd, $J_{2,1}$ = 2.8 Hz, $J_{2,3}$ = 11.7 Hz, 1H, H-2), 3.02 (m, H-4'a,), 2.84 (m, H-4'b), 2.15 (s, 3H, $COCH_3$), 2.03 (s, 3H, $COCH_3$), 2.00 (s, 3H, $COCH_3$); **^{13}C NMR** (125 MHz, CD_3OD) δ: 173.1 (C_q, CO), 172.1 ($2C_q$, CO), 171.9 (C_q, CO), 171.5 (C_q, CO), 167.1 (C_q, CO), 157.0 (C_q, CO), 97.6 (C-1), 97.1 (C_q), 70.1 (C-5), 68.7 (C-4), 67.2 (C-3), 62.8 (C-6), 53.4 (C-5'), 42.3 (CH_2, Gly), 37.4 (C-2), 31.8 (C-4'), 20.5-20.4 ($3CH_3$, CH_3).

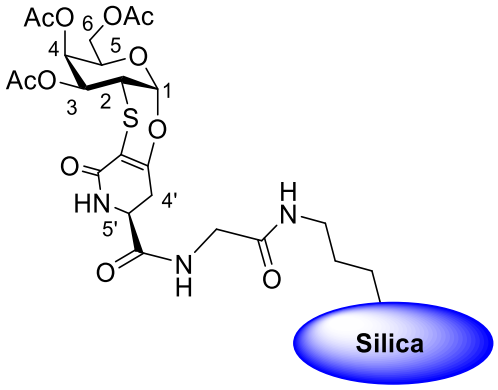
Synthesis of compound 40



Compound **27** (140 mg, 0.31 mmol) was suspended in dry CH_2Cl_2 (3 mL), HBTU (174 mg, 0.46 mmol), DIPEA (89 µL, 0.93 mmol) and glycine benzyl ester hydrochloride (74.6 mg, 0.37 mmol) were added and the suspension was stirred at room temperature for 4 h. After complete conversion, the mixture was diluted with CH_2Cl_2 and washed with HCl 1M (x3). The organic layer was dried over Na_2SO_4 and concentrated under vacuum. The crude was purified by flash chromatography (CH_2Cl_2/CH_3OH 95:5) to give **40** as white solid (168 mg, 91%).

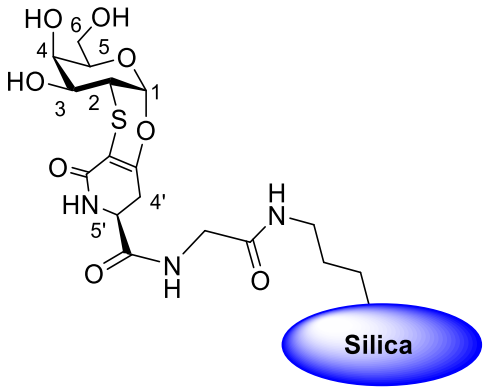
$[\alpha]^{26}_D$ = +126.5 (c 0.001, CH_3OD); **ESI-MS *m/z* (%):** 629.33 (100) $[M+Na]^+$, 1235.08 (97) $[2M+Na]^+$; **^{1}H NMR** (500 MHz, $CDCl_3$) δ: 7.39-7.30 (m, 5H, Bn), 6.98 (t, 1H, $J_{NH,CH2Gly}$ 5.6 Hz, NH), 6.56 (d, 1H, $J_{NH,H\text{-}5'}$ 3.2 Hz, NH), 5.69 (d, $J_{1,2}$ = 2.7 Hz, 1H, H-1), 5.42 (dd, $J_{4,3}$ = 3.1 Hz, $J_{4,5}$ = 1.1 Hz, 1H, H-4), 5.16-5.15 (m, 2H, CH_2Ph), 5.02 (dd, $J_{3,2}$ = 11.7 Hz, $J_{3,4}$ = 3.1 Hz, 1H, H-3), 4.46-4.41 (m, 1H, H-5), 4.24-4.20 (m, 1H, H-5'), 4.17-4.13 (m, 2H, H-6a, H-6b), 4.12-4.08 (m, 2H, CH_2, Gly), 3.60 (dd, $J_{2,1}$ = 2.7 Hz, $J_{2,3}$ = 11.7 Hz, 1H, H-2), 3.01-2.88 (m, H-4'a, H-4'b), 2.16 (s, 3H, $COCH_3$), 2.07 (s, 3H, $COCH_3$), 2.02 (s, 3H, $COCH_3$); **^{13}C NMR** (125 MHz, $CDCl_3$) δ: 170.6 (C_q, CO), 170.2 (C_q, CO), 170.1 (C_q, CO), 170.0 (C_q, CO), 169.5 (C_q, CO), 164.8 (C_q, CO), 155.0 (C_q), 135.22 (C_q, Ph), 128.8 (CH, Ph), 128.7 (C_q, Ph), 128.5 (CH, Ph), 97.3 (C_q), 96.0 (C-1), 69.1 (C-5), 67.6 (CH_2Ph), 67.2 (C-4), 66.1 (C-3), 61.7 (C-6), 52.4 (C-5'), 41.6 (CH_2, Gly), 36.6 (C-2), 30.9 (C-4'), 20.8-20.7 ($3CH_3$, CH_3).

Synthesis of compound 41



To a solution of **39** (90 mg, 0.17 mmol) in DMF (2 mL), HBTU (197 mg, 0.52 mmol), DIPEA (121 µL, 0.70 mmol) and aminopropyl-functionalized silica gel (88 mg, 1 mmol/g NH_2 loading) were added. The suspension was stirred at room temperature for 24 h, then was filtered and washed several times with DMF (x5) and CH_2Cl_2 (x5) to give **41** (110 mg).

Synthesis of compound 42



A suspension of **41** (75 mg) in NH_3 in CH_3OH 4 M (1.5 mL) was stirred at room temperature for 48 h, then was filtered and washed several times with CH_3OH (x5) and CH_2Cl_2 (x5) to give **42** (60 mg).

Part III

A TFThr mimetic

CHAPTER

9

Interaction of a TFThr tumor-associated antigen mimetic with Gal-3

9.1 Introduction

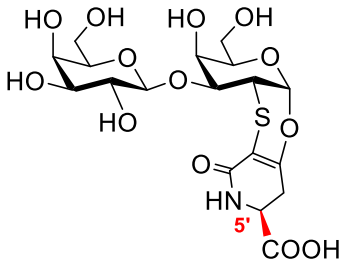
Encouraged by the promising results obtained with the simplest TnThr antigen, the glycomimetic strategy was applied for the synthesis of a TFThr mimetic. Indeed, the crystal structure of Gal-3 CRD complexed with the TF antigen showed the TF antigen as an efficient ligand for Gal-3.[169] As already pointed out (**Chapter 3 - Section 3.2.2**) TF antigen, a "*pan-carcinoma antigen*", overexpressed in approximately 90% of adenocarcinomas[2], plays a dramatic role in cell-to-cell adhesion and cancer cell proliferation. In cancer spread, the interaction of TF with specific lectins such as galectins is crucial. [61,170] Galectins are multivalent, nonenzymatic carbohydrate binding proteins, characterized by a conserved carbohydrate-recognition domain (CRD) and a high affinity and specificity for β-galactose-containing oligosaccharides. [171] Indeed, it has

been shown that also Gal-3, one of the most studied galectins, is overexpressed in malignant tumor cells and induces MUC1 clustering on the tumor cell surface by binding to the TF antigen.[172] As a consequence of this interaction, tumor cells invade vessels and survive in the circulation, avoiding anoikis.[173] The interruption of this interplay with a physiologically stable TF mimetic, competing with the native one for the same binding pocket, could prevent MUC1 clustering and reduce metastasis. For this reason, the binding properties of the TF mimetic, vs. the purified Gal-3 CRD, were studied by combining NMR experiments with X-ray crystallography and isothermal titration calorimetry (ITC) assays.[174]

9.2 Results and Discussion

9.2.1 Synthesis of a TFThr mimetic

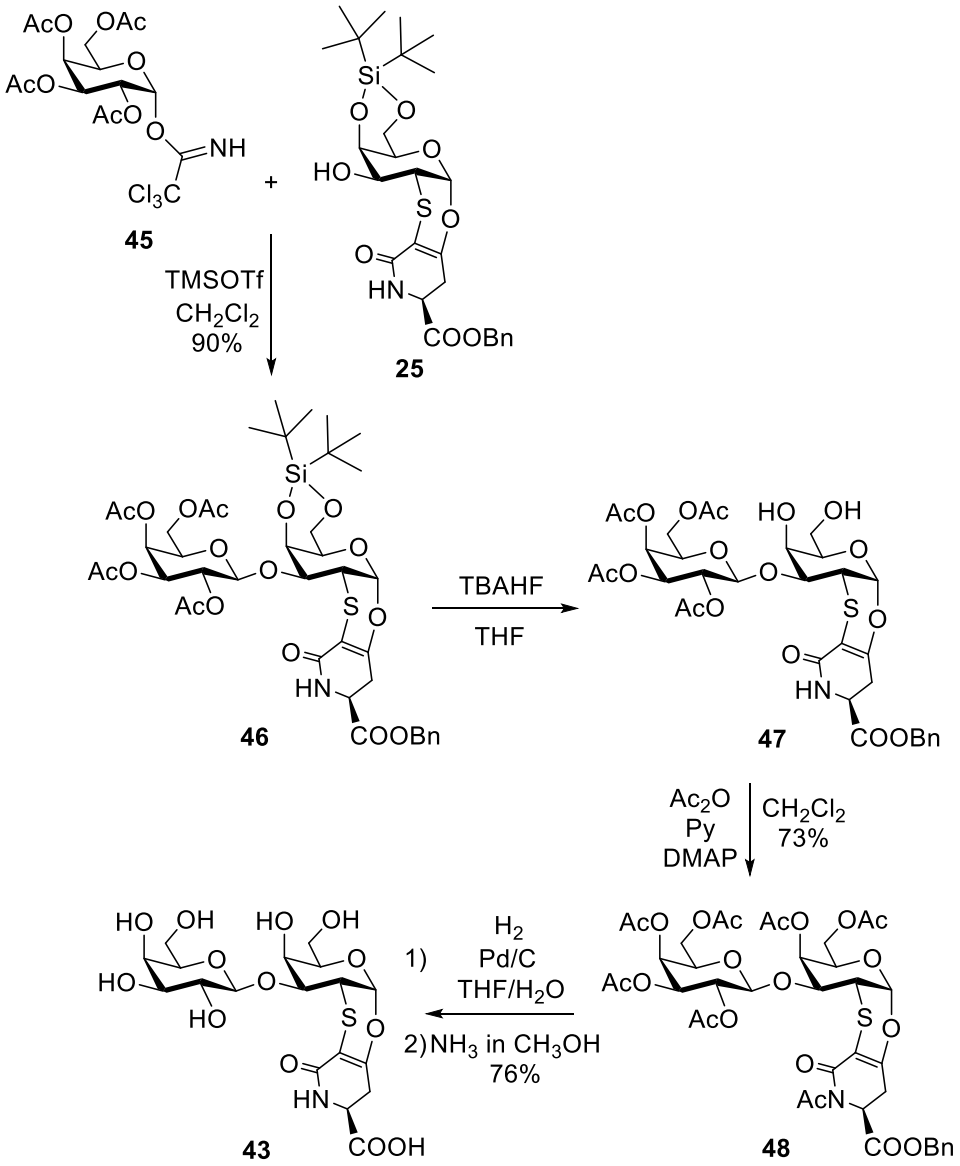
In 2009, a first attempt to synthesize the TFThr mimetic **43** was reported by Nativi *et al*. resulting in a mixture of diastereoisomers due to an isomerization at the lactam ring level (**Figure 9.2.1.1**).[82]



43

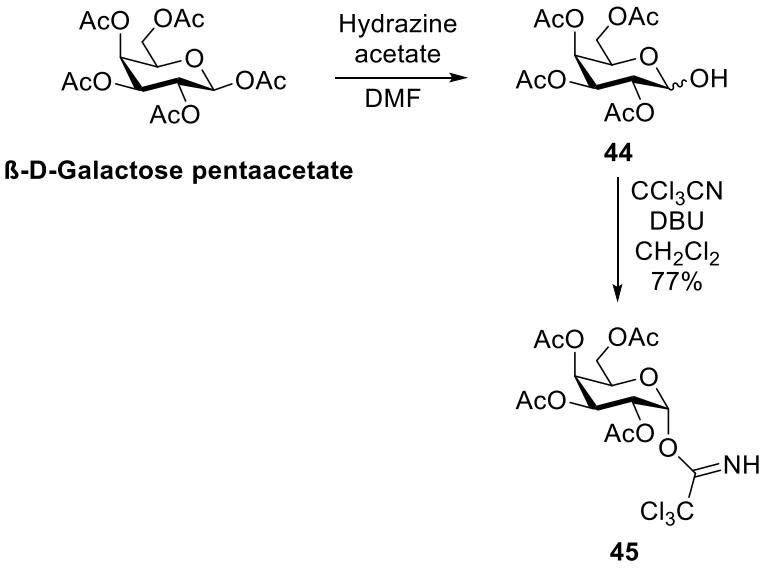
Figure 9.2.1.1 Structure of TFThr mimetic **43**. The stereocenter object of racemization is highlighted.

Capitalizing on this misstep, starting from the protected TnThr mimetic **25** by the formation of a β1-3 bond with a Gal residue, the TFThr mimetic was efficiently isolated (**Scheme 9.2.1.1**).



Scheme 9.2.1.1 Synthesis of the TFThr mimetic **43**

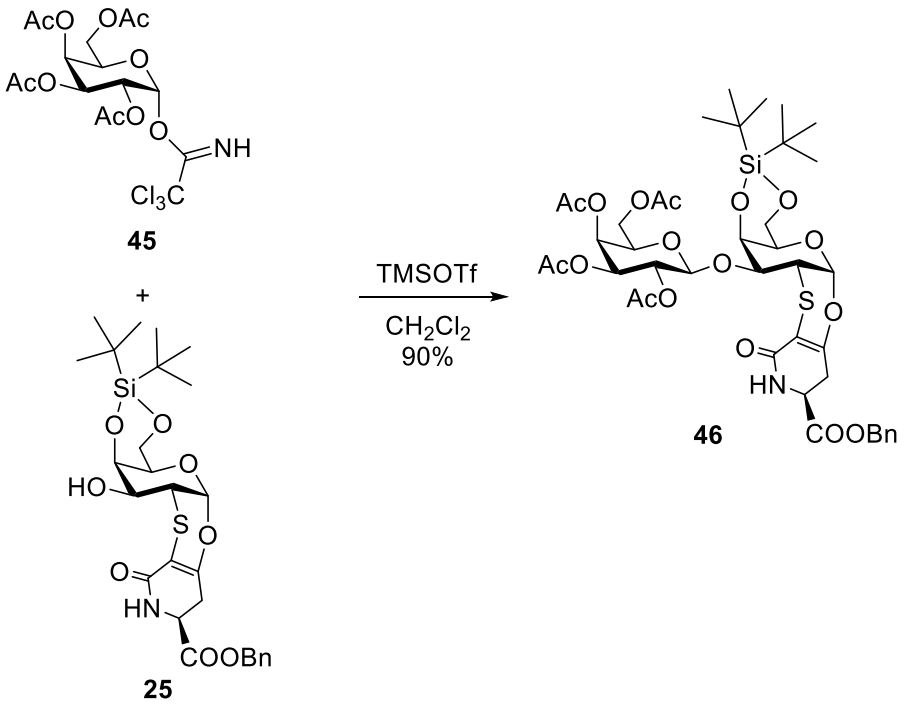
In the first steps, starting from the β-D-galactose pentaacetate, after a selective anomeric deprotection with hydrazine acetate, the anomeric unprotected derivative **44** was isolated and activated as trichloroacetimidate to give pure **45** in a 77% yield over 2 steps. (**Scheme 9.2.1.2**).



Scheme 9.2.1.2 Synthesis of the glycosyl donor **45**

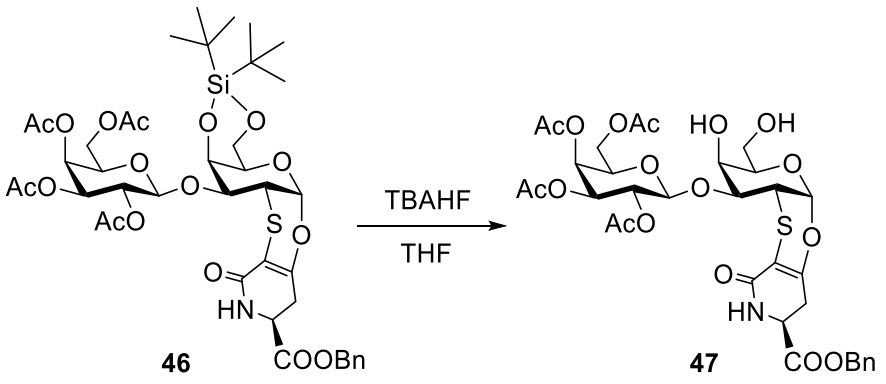
An excess of trichloroacetimidate donor **45** was reacted under Schmidt conditions[175] with the protected TnThr mimetic acceptor **25**, bearing a free hydroxyl group at C-3, to form the fully protected disaccharide **46** as a single β-diastereoisomer in a 90% yield. (**Scheme 9.2.1.3**).

From the deprotection of silylidene (**Scheme 9.2.1.4**) with a freshly prepared solution of TBAHF, crude **47** was isolated and used as crude in the following reaction.



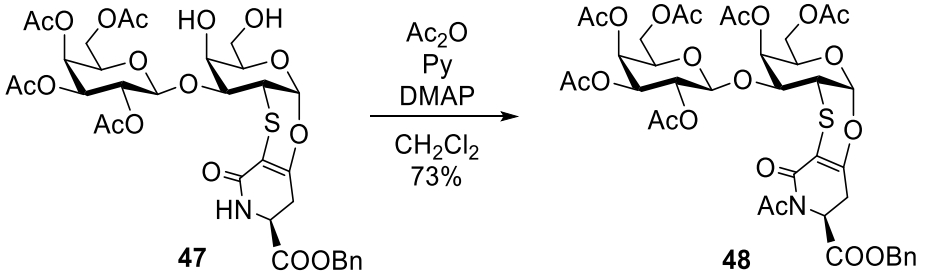
Scheme 9.2.1.3 Synthesis of the protected TFThr mimetic **46**

Subsequently, for purification reasons, the peracetylated derivative **48** was prepared (73% yield over 2 steps) (**Scheme 9.2.1.5**), upon treatment of **47** with Ac_2O, pyridine and a catalytic amount of DMAP. Noteworthy,



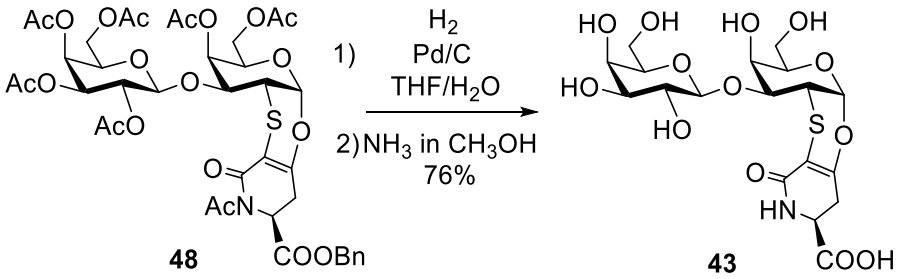
Scheme 9.2.1.4 Removal of silylidene protecting group and synthesis of **47**

prolonged reaction time, necessary for a complete acetylation of the starting material, caused the undesired protection of the lactam nitrogen.



Scheme 9.2.1.5 Synthesis of the peracetylated TFThr mimetic **48**

Finally, in two consecutive steps, (**Scheme 9.2.1.6**) the benzyl group was removed by palladium catalyzed hydrogenation and the acetyls were cleaved by treatment with a solution of NH_3 in CH_3OH (2M). We thus isolated the complete deprotected TFThr mimetic as a single diastereoisomer in a 76% yield.



Scheme 9.2.1.6 Deprotection of the derivative **48** and synthesis of the TFThr mimetic **43**.

9.2.2 NMR and ITC binding studies of 43 vs. Gal-3 CRD

First of all, for the evaluation of the molecular recognition by NMR, a ^{15}N-labelled Gal-3 CRD, expressed and purified as previously reported[176] was titrated with increasing concentrations of **43** (**Figure 9.2.2.1**).

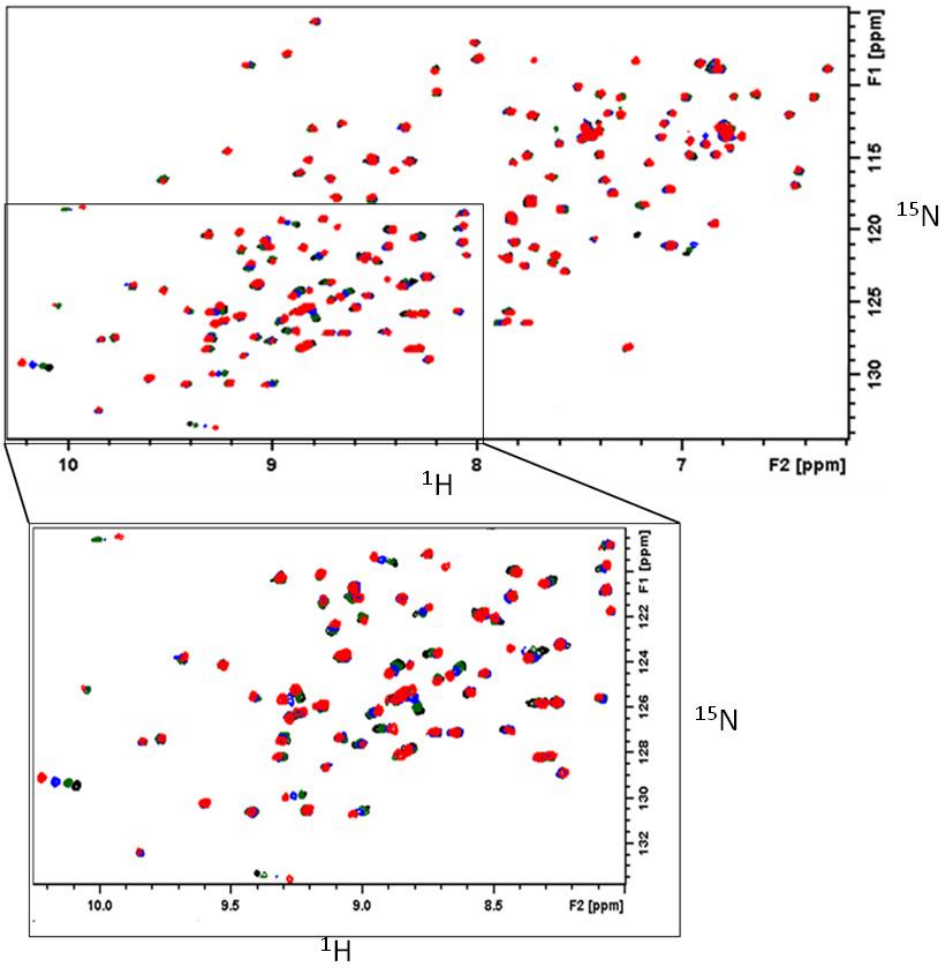


Figure 9.2.2.1 Overlap of the ^{1}H,^{15}N-HSQC (full spectra and an expansion of the most perturbed region) of Gal-3 CRD in the presence of **43** at distinct L/P molar ratios. Black spectrum = L/P 1:0, green spectrum = L/P 1:1, blue spectrum = L/P 1:8, and red spectrum L/P 1:50.

A $^{1}H–^{15}N$ heteronuclear single-quantum coherence ($^{1}H–^{15}N$ HSQC), chemical shift perturbation analysis of the cross-peaks corresponding to the amino acids of Gal-3 was performed. The associated combined chemical shift perturbation (*Δδcomb*) was estimated considering the variation of chemical shifts both in the $^{1}H–^{15}N$ dimensions[177] and a cut-off line was calculated to better determine the affected and unaffected residues of Gal-3 CRD.[177]

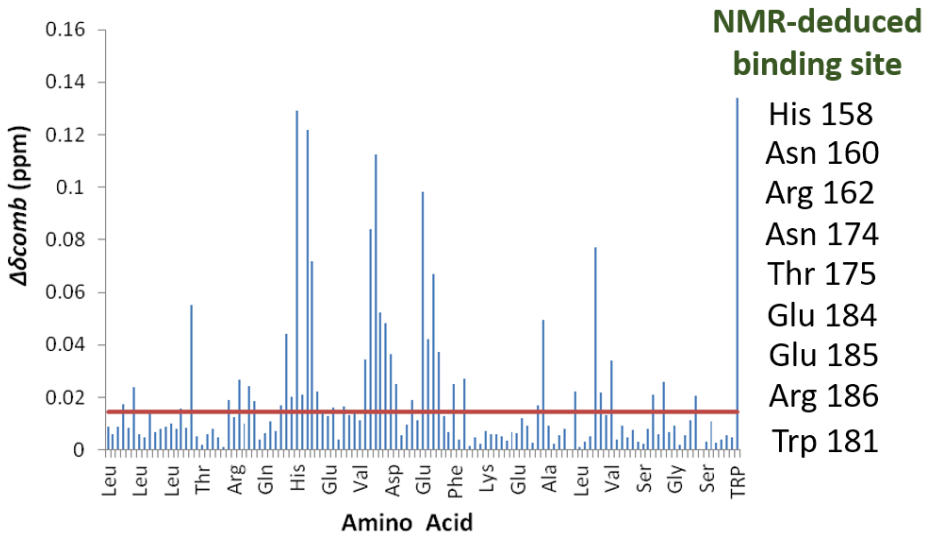


Figure 9.2.2.2 Plot of the $^{1}H,^{15}N$ combined chemical shift (*Δδcomb*) for each amino acid of Gal-3 CRD considering 50 equivalents of **43**. The orange line corresponds to the value of cut-off that dictates the amino acids of Gal-3 CRD that are mostly perturbed (NMR-deduced binding site).

The chemical shift perturbation analysis unequivocally indicated a confined Gal-3 binding site for the TFThr mimetic **43** (**Figure 9.2.2.2**). The most perturbed residues upon addition of **43** (His158, Asn160, Arg162, Asn174, Thr175, Glu184, Glu185, Arg186 and the indole NH of Trp181)

constitute the NMR-deduced binding site, which is mostly coincident to that reported for the natural TF antigen and which has been previously fully characterized by X-ray crystallography and NMR studies.[169,178,179]

The binding was also monitored by analyzing the line broadening of the **43** ^{1}HNMR signals with increasing concentration of Gal-3 CRD. (**Figure 9.2.2.3**).

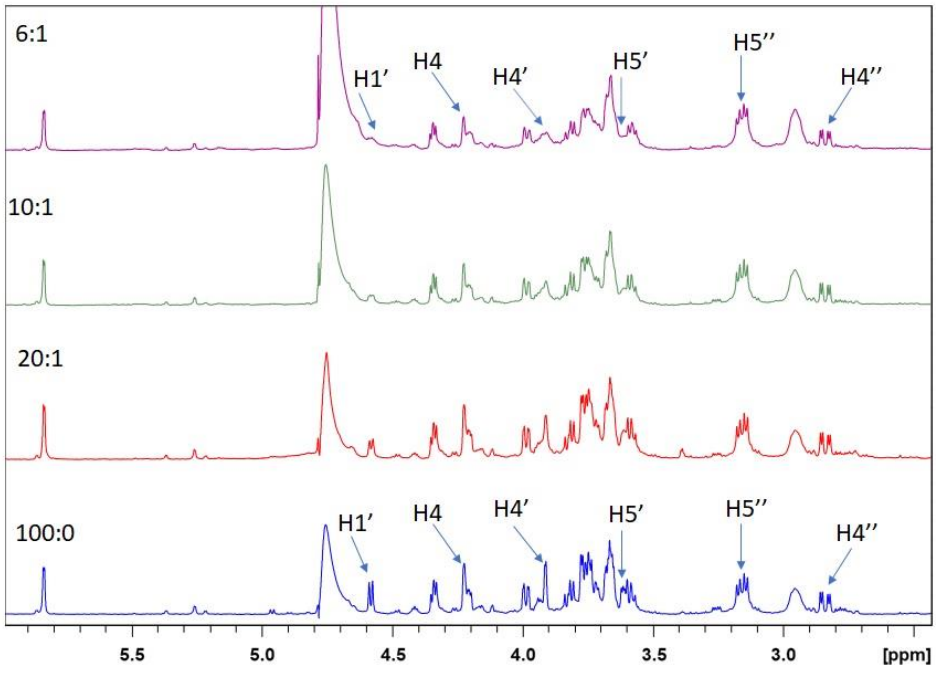


Figure 9.2.2.3 Line broadening of **43** ^{1}HNMR signals function of different concentrations of the Gal-3 CRD.

The H4' and H1' signals of the non-reducing Gal moiety exhibited a significant broadening of their line widths. The other protons of the Gal moiety together with some GalNAc signals also suffered from a moderate decreasing in signal intensity. In contrast, the line widths of the lactam ring protons were mostly unperturbed in the presence of Gal-3 CRD. This

result indicates that the disaccharide interacts directly with Gal-3 CRD, while the tricycle moiety is exposed to the bulk solvent. (**Figure 9.2.2.4**). This observation also indicates that **43** presents a binding epitope similar to that reported for the native antigen.[178]

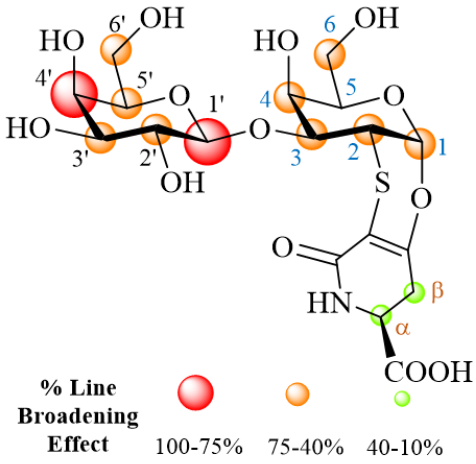


Figure 9.2.2.4 Line broadening effect derived epitope map of the **43** in the presence of Gal-3 CRD.

Moreover, the observed effects on the chemical shifts were used to calculate the dissociation constants (K_D) as function of the concentration of **43**.[180] In particular, three residues were selected, based on their direct participation to engage the natural TF antigen through hydrogen bonding, Asn160 with the Gal unit, Glu184 with the GalNAc moiety, and Trp181 with CH–π stacking interactions with the α face of the galactose ring.[169]

The results yielded similar K_D values of 208, 188, and 233 μm for Asn160, Glu184, and the lateral chain of Trp181, respectively (**Figure 9.2.2.5**). The average of the K_D yielded an apparent K_D (K_{Dapp}) of 210 μm. This result is

in the same range as those reported by isothermal titration calorimetry (ITC) for the natural TF (248 µm) and natural TFThr (272 µm).[179]

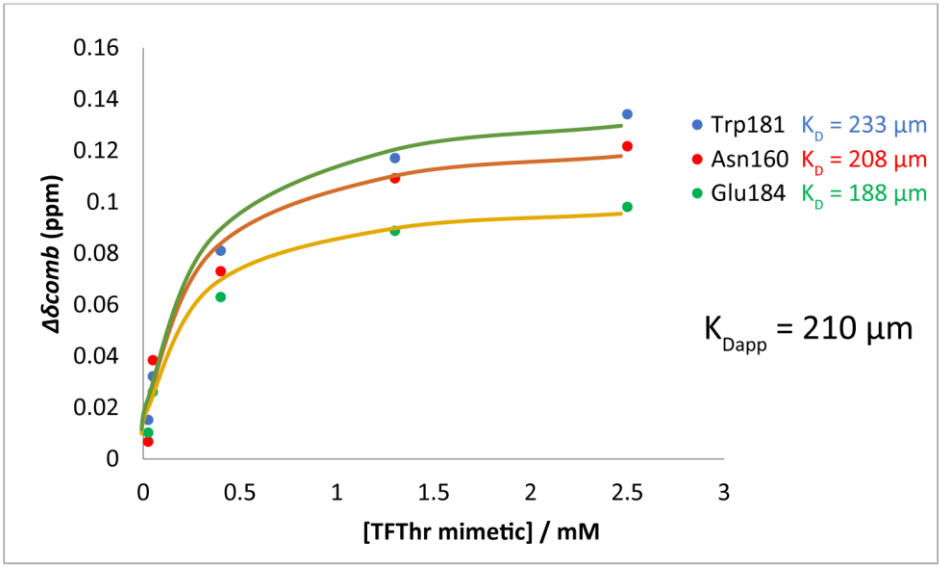


Figure 9.2.2.5 Plot of the *Δδcomb* (ppm) for three amide signals (backbone and lateral chain). The lines correspond to the fittings to one binding site model.

Therefore, ITC measurements were performed to confirm the dissociation value obtained via NMR. A slightly more favorable change in Gibbs free energy (ΔG) associated with the binding of **43** to Gal-3 than that reported for TF and TFThr was observed. Thermodynamic data reflect a slight improvement in the estimated K_D value for the TF mimetic **43** (166 µm) confirming the NMR data. Upon inspection of the delicate enthalpy/entropy balance, thermodynamics parameters for TFThr mimetic **43** were compared with those previously reported by Cudic *et al.* for natural TF and natural TFThr counterparts (**Figure 9.2.2.6**).[178] While the binding event of natural derivatives event is enthalpically driven (13–14 kcal mol^{-1}) the enthalpy gain is only 4 kcal mol^{-1} for the TnThr mimetic

43. This striking loss in the enthalpic contribution (8–10 kcal mol^{-1}) can, in principle, be related to the loss of some of the interactions. For instance, since the NHAc of the GalNAc moiety is absent in **43**, the interactions established through this residue cannot take place.[169] However, a compensation in the entropic contribution was detected for the **43** binding to Gal-3 CRD, which balances the overall ΔG. The introduction of the lactam ring at the C1–C2 region confers limited conformational flexibility at the reducing end of **43** that may be the reason of the reduced entropic penalty. Although these facts could partially support some of the enthalpy loss and entropy gain observed in the binding, they do not completely explain the observed behavior.

Ligand	K_D (μM)	ΔG (kcal/mol)	ΔH (kcal/mol)	-TΔS (kcal/mol)	n
TFThr mimetic	166	-5.15	-4.27±0.69	-0.88	1
TF	248	-4.94	-14.05±0.07	9.11	0.93
TFThr	272	-4.88	-12.63±0.43	7.75	0.90

Figure 9.2.2.5 Thermodynamics parameters estimated by ITC for the binding of TFThr mimetic **43** to Gal-3 CRD and data for natural TF and TFThr extracted from the work of Cudic.[178]

9.2.3 X-ray of the Gal-3 CRD-TFThr mimetic complex

Despite, the NMR analysis point out a very similar binding mode (binding epitope and bioactive conformation) for the TFThr mimetic **43** and the natural TF antigen, the data did not afford a clear-cut explanation of the observed enthalpy/entropy compensation phenomenon. For this reason, X-ray crystallographic analysis of the complex was also carried out. The crystal structure of the complex between **43** and Gal-3 CRD was

determined by molecular replacement methods using the unliganded Gal-3 CRD structure (PDB ID : 3ZSL). The 3D structure of Gal-3/TFThr mimetic **43** was solved to 1.1 Å resolution (PDB ID : 6G0V) and resulted very similar to that reported for the Gal-3 CRD/TFThr complex (PDB ID : 3AYA).[169]

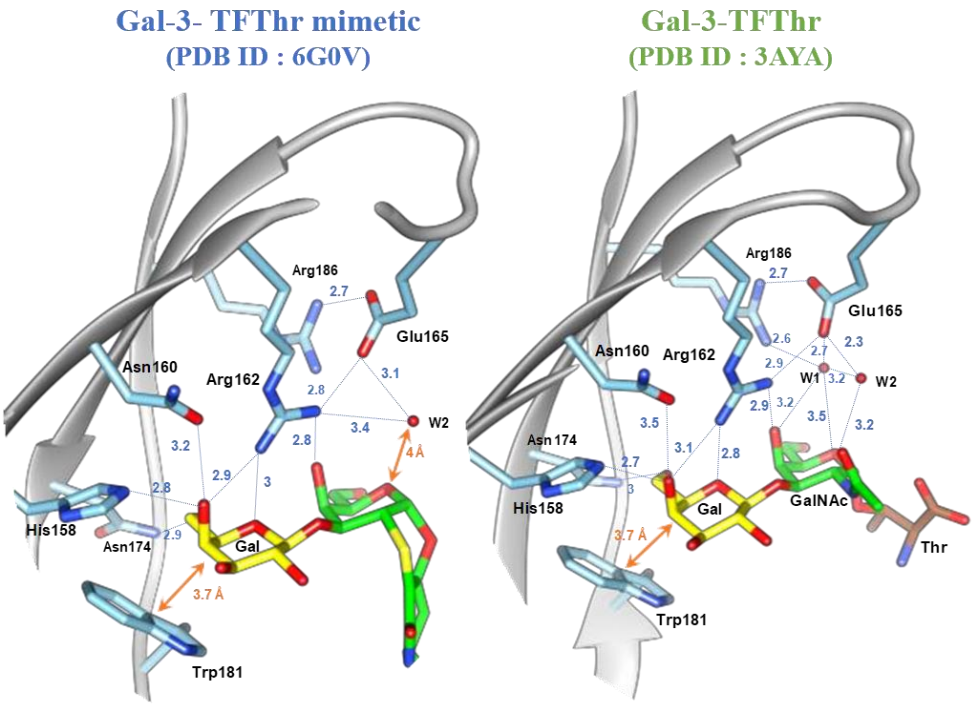


Figure 9.2.2.2 Side-by-side view of the binding site of Gal-3 CRD domain complexed with TFThr mimetic **43** (left) and natural TFThr antigen (right). Hydrogen bonds = dashed blue lines, CH– π, as well as the distance between W2 and O3 of TFThr mimetic = orange arrows. Distances are in ångströms.

Regarding the bound state, the *2mF$_o$-DF$_c$* electron density map showed that, as expected, both carbohydrate rings of the **43** present a $^4C_1$ chair conformation with a slightly distorted *exo*-anomeric orientation for the glycosidic linkage (H1'-C1'-O-C3 ≈ 35°), similar to that observed for the

natural TF antigen. Furthermore, even crystal structure observations indicate that the lactam ring of **43** is exposed to the solvent and does not establish any contacts with the protein surface. Moreover, Gal-3 CRD/TFThr mimetic complex structure unveils that Gal-3 conformationally selects the major conformer of **43** present in solution, previously proposed using NMR and molecular dynamics (MD) simulations.[82] As already deduced by NMR, the Gal moiety of **43** interacts with the side chains of residues His158, Asn160, Arg162, and Asn174 through hydrogen bonds, while the side chain of Trp181 displays CH–π stacking interactions to the α-face of Gal (**Figure 9.2.3.1**). However, important differences in the hydrogen bonding networks were found around the reducing moieties in the two complexes. In the Gal-3 CRD/TFThr crystal structure, there is a unique hydrogen bond network through water molecules that involves the Arg186 guanidinium group and the Glu165 carboxylate group of Gal-3 with the endocyclic oxygen atom and the carbonyl group of the GalNAc moiety which is absent in the crystal structure of the complex with **43**. The lack of this particular network could explain enthalpic loss measured by ITC. As previously discussed for **43**, the strength of binding and affinity to Gal-3 is driven by an entropic contribution. In addition to the limited conformational flexibility, the release of water molecules to the bulk in the complex of **43** should likely play a role in the entropy gain. Indeed, other examples of glycan–protein interactions were shown to have a favored entropic effect attributed to the release of bound water molecules to the bulk.[181] In this context, the chemical nature of the TF mimetic **43**, in particular the introduction of the lactam ring, that confers an amphiphilic character, and the associated changes in its solvation

features between the free and bound states should also influence the entropic term. Thus, it is likely that in this case as well, water is behind the observed entropy gain, which strikingly even surpasses the large loss in enthalpy.

9.3 Conclusions

In summary, in this chapter a straightforward synthesis of the diastereoisomeric pure TFThr mimetic **43** and binding studies with Gal-3 CRD were reported**.** A combination of NMR, ITC and X-ray crystallography was employed to unveil the molecular recognition's nature. The obtained NMR data strongly suggest that the chemical modification at the reducing unit of the TFThr mimetic **43** does not affect the global affinity for Gal-3 and unequivocally indicates a confined Gal-3 binding site for the **43** with an NMR-deduced binding site, which is mostly coincident to that reported for the natural TF antigen. The slight increase in the NMR calculated dissociation constants (K_{Dapp}) for **43** was confirmed by ITC. In the analysis of ITC data, in this particular case the enthalpy loss is compensated by an entropy gain, showing how this significant entropy change can be achieved by a simple modification of the chemical nature of the ligand, keeping its basic binding features but generating an amphiphilic surface. Although, the nature of this phenomena is not completely understood, the present example illustrates the possibility of improving the entropy term by designing a rigid, amphiphilic artificial ligand that interacts at the same binding site of the natural counterpart. Furthermore, X-ray analysis confirmed the data obtained by NMR and ITC indicating that the same binding epitope and conformation at the bound state as those of the

natural analogue were retained. Hence, in the future this new glycomimetic could be exploited for multivalent presentations, as already done with the TnThr mimetic, to interfere Gal-3 aberrant interactions in cancer progression and metastasis.

9.4 Experimental section

9.4.1 General methods

The combined chemical shift calculation and determination of the K_D from $^{1}H,^{15}N$-HSQC titration analysis: For the evaluation of the behavior of individual amino acids upon the addition of **43** the combined amide proton and nitrogen chemical shift differences were determined (Δδcomb) using equation **1**:

$$\Delta\delta comb = \sqrt{(\Delta\delta H)^2 + (\frac{\gamma_N}{\gamma_H}\Delta\delta N)^2} \qquad \text{eqn (1)}$$

where $\Delta\delta H$ is the difference between the chemical shift of one of the concentrations of the titration in the presence of **43** and the chemical shift of the proton in the spectrum without ligand; $\Delta\delta N$ is the difference between the chemical shift of the signal obtained in one of the concentrations of the titration in the presence of **43** and the chemical shift of the nitrogen atom in the spectrum without the ligand and the $\frac{\gamma_N}{\gamma_H}$ is the quotient between the gyromagnetic constant of the nitrogen and the gyromagnetic constant of the proton. With the values of $\Delta\delta comb$ calculated for the individual amino acid a column chart using the highest value of **43** concentration was built. To distinguish the residues with more participation in the interaction process, the standard deviation to zero (SD) of the $\Delta\delta comb$ was determined to find a reasonable cut-off criterion.[177] This is an iterative process, where the standard deviation and cut-off are recalculated every time a residue has a value of $\Delta\delta comb$ higher than the cut-off. These residues are removed from the calculation, until

all values of $\Delta\delta comb$ are below the cut-off. The combined chemical shifts of the NH group resonance from polypeptide assignment of Asn160, Glu184 and the indole NH of Trp181 amino acids of Gal-3 CRD were used to obtain the dissociation constant (K_D) from the titration experiment according to equation **2**:[180]

$$\Delta\delta comb = \Delta\delta max \frac{(K_D+[L]_0+[R]_0)-\sqrt{(K_D+[L]_0+[R]_0)^2-4\times([L]_0\times[R]_0)}}{2\times[R]_0} \quad \text{eqn (2)}$$

where $[L]_0$ is the ligand concentration, $[R]_0$ is the protein concentration, $\Delta\delta comb$ is the combined chemical shift calculated by eqn **1** and the $\Delta\delta max$ is the maximum chemical shift for each amino acid between the free and bound state.

X-Ray crystallography: Data collection and refinement statistics: Statistics of data collection, processing and refinement. Statistics for the highest-resolution shell are shown in parentheses.

Data collection and processing	
Space group	$P2_1\,2_1\,2_1$
Unit-cell parameters (Å)	a = 35.9, b = 58.2, c = 62.8
Resolution range (Å)	42.67 - 1.10 (1.14 - 1.10)
Matthews coefficient, V_M	2.34
Solvent content (%)	47.5
Protein molecules per asymmetric unit	1
Mosaicity (°)	0.2
I/σ(I)	17.8 (2.3)
Wilson B-factor	11.9
R_{merge}†	0.040 (0.506)
$R_{p.i.m.}$+	0.019 (0.262)

Half-dataset correlation CC1/2	0.999 (0.834)
Multiplicity	5.2 (4.0)
Total reflections	217687 (2863)
Unique reflections	41602 (711)
Completeness (%)	76.2 (29.8)
Refinement statistics	
Protein atoms	1192
Ligand atoms	34
Solvent molecules	223
R_{work}‡	0.153 (0.237)
R_{free}§	0.177 (0.238)
R.m.s.d. bond lengths (Å)	0.007
R.m.s.d. bond angles (°)	0.97
Average B-factor (Å^2)	17.6
Protein	15.1
Main chain	13.0
Side chain	18.4
TFThr mimetic	31.6
Water molecules	29.1
Ramachandran plot	
Residues in favored regions (%)	97.1
Residues in allowed regions (%)	2.9
Outliers (%)	0.0
PDB accession code	6G0V

† $R_{merge} = \frac{\sum_{hkl}\sum_{i=1}^{n}|I_i(hkl)-\bar{I}(hkl)|}{\sum_{hkl}\sum_{i=1}^{n}I_i(hkl)}$, where I is the observed intensity, and $\bar{I}$ is the statistically-weighted average intensity of multiple observations. ‡

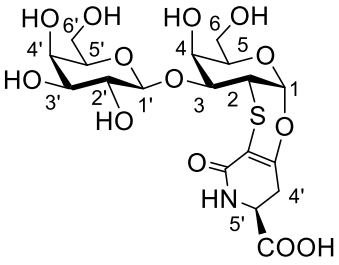
$R_{p.i.m.} = \frac{\sum_{hkl}\sqrt{1/(n-1)}\sum_{i=1}^{n}|I_i(hkl)-\bar{I}(hkl)|}{\sum_{hkl}\sum_{i=1}^{n} I_i(hkl)}$, a redundancy-independent version of R_{merge}. ‡ $R_{work} = \frac{\sum_{hkl}||F_{obs}(hkl)|-|F_{calc}(hkl)||}{\sum_{hkl}|F_{obs}(hkl)|}$, where $|F_{calc}|$ and $|F_{obs}|$ are the calculated and observed structure factor amplitudes, respectively. § R_{free} is calculated for a randomly chosen 5% of the reflections.

9.4.2 Synthesis

General method

Analytical grade solvents and commercially available reagents were used without further purification. For anhydrous reactions, solvent stored over 3 Å molecular sieves were used. Silica gel flash column chromatography purifications were performed using Geduran® Si 60 (0.040-0.063 mm) or using the Biotage Isolera system and SNAP silica cartridges. TLC analyses were performed on glass Merck silica gel 60 F_{254} plates. ^{1}H NMR, ^{13}C NMR and 2D-NMR spectra were recorded on a 500 MHz Bruker AVANCE II at 298 K, unless otherwise stated. All chemical shifts are reported in parts per million (δ) referenced to residual nondeuterated solvent. Multiplicity abbreviation: b = broad, s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet were used. ESI-MS spectra were carried out on a linear ion-trap double quadrupole mass spectrometer using electrospray ionization (ESI) technique (LTQ-XL - Thermo Fisher). Optical rotation measurements were performed on a JASCO DIP-370 polarimeter. Melting point were recorded on a STUART SMP3 version 2.0 or a BUCHI 510.

Synthesis of compound 43

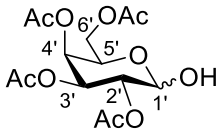


The peracetylated derivative **48** (150 mg, 0.17 mmol) was suspended in a mixture of THF/H_2O 10:1 (3 mL); then Pd/C (10 wt.%, 10 mg) was added and the reaction was stirred at room temperature overnight, under H_2 atmosphere. After reaction completion, the mixture was diluted with EtOAc and filtered through a pad of Celite®. The filtrate was concentrated to dryness and suspended in 2 M solution of NH_3 in CH_3OH (2 mL). After stirring for 6 h at room temperature the solvent was evaporated under vacuum to give pure **43** (64 mg, 76%)

$[\alpha]^{29}_D$ = +1.7 (c 0.001, H_2O); **ESI-MS *m/z* (%):** 494.08 (100) $[M-H]^-$; **^{1}H NMR** (500 MHz, D_2O) δ: 5.75 (d, $J_{1,2}$ = 2.6 Hz, 1H, H-1), 4.50 (d, $J_{1'-2'}$ = 8.0 Hz, 1H, H-1'), 4.25-4.29 (X part of an ABX system $J_{X,A}$ = 7.2 Hz, $J_{X,B}$ = 6.8 Hz, 1H, H-5''), 4.14 (m, 1H, H-4), 4.12 (bs, 1H, H-5'), 3.89-3.85 (dd, $J_{3,4}$ = 2.5 Hz, $J_{3,2}$ = 11.0 Hz, 1H, H-3), 3.66-3.68 (dd, $J_{3'-4'}$ = 3.4 Hz, $J_{3'-2'}$ = 10.6 Hz, 1H, H-3'), 3.77-3.60 (A part of an ABX system, $J_{A,B}$ = 8.1 Hz, 1H, H-6'a), 3.74-3.72 (A part of an ABX system, $J_{A,B}$ = 11.4 Hz, 1H, H-6a), 3.71-3.66 (B part of an ABX system, $J_{B,A}$ = 12.1 Hz, 1H, H-6b), 3.90 (d, $J_{4'-3'}$ = 3.5 Hz, 1H, H-4'), 3.61-3.67 (B part of an ABX system, $J_{B,A}$ = 8.1 Hz, 1H, H-6'b), 3.59-3.56 (dd, $J_{2,1}$ = 2.7 Hz, $J_{2,3}$ = 10.7 Hz, 1H, H-2), 3.54-3.50 (dd, $J_{2'-1'}$ = 7.8 Hz, $J_{2'-3'}$ = 10.4, 1H, H-2'), 3.52 (bs, 1H, H-5), 3.11-3.03 (A part of an ABX system, $J_{A,B}$ = 17.0 Hz, $J_{A,X}$ = 7.8 Hz, 1H, H-4''a), 2.77-2.71 (A part of an ABX system, $J_{B,A}$ = 17.1 Hz, $J_{B,X}$ = 7.8 Hz, 1H, H-4''b). **^{13}C NMR** (125 MHz, D_2O) δ: 170.38 (CO), 165.8 (CO), 156.2 (C_q), 103.5 (C-1'), 96.5 (C_q), 96.3 (C-1), 77.7 (C-3), 72.6 (C-4),

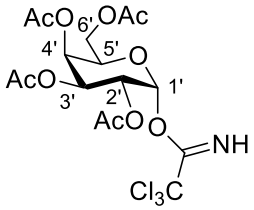
70.5 (C-5), 70.7 (C-5'), 69.8 (C-3'), 68.9 (C-4'), 67.13 (C-2'), 67.08 (C-6'), 66.9 (C-6), 60.5 (C-5''), 37.1 (C-2), 31.4 (C-4'').

Synthesis of compound 44



To a solution of β-D-galactose pentaacetate (5.00 g, 12.81 mmol) in DMF (18 mL), hydrazine acetate (1.76 g, 19.20 mmol) was added and the suspension was stirred at room temperature. After 2 h, the mixture was diluted with EtOAc and washed with H_2O (x5). The organic layer was dried over Na_2SO_4, concentrated under vacuum to give crude **44** (6.14 g) that was used without further purification.

Synthesis of compound 45

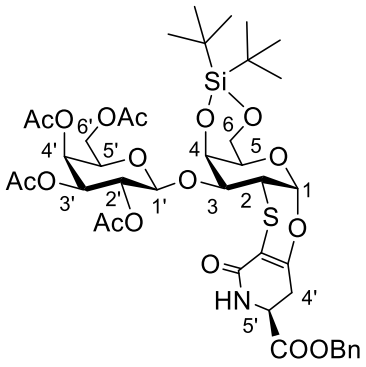


A solution of crude 2,3,4,6-tetra-O-acetyl-D-galactopyranoside **44** (6.14 g, 12.81 mmol) in dry CH_2Cl_2 (90 mL) was cooled to - 0 °C and DBU (1.96 g, 12.81 mmol) and CCl_3CN (37.00 g, 256.25 mmol) were added. The dark brown solution was stirred at room temperature under N_2 atmosphere. After 5 h, the reaction was diluted with CH_2Cl_2 and washed with NH_4Cl (x2) and H_2O (x3). The organic layer was dried over Na_2SO_4 and concentrated under vacuum. The crude was purified by flash chromatography (EP/EtOAc 2:1 + Et_3N 0.1%) to give pure **45** as yellowish solid (4.86 g, 77%).

^{1}H NMR (500 MHz, $CDCl_3$) δ: 8.66 (bs, 1H, NH), 6.63 (d, $J_{1',2'}$ = 3.4 Hz, 1H, H-1'), 5.59-5.57 (dd, $J_{4',3'}$ = 3.2 Hz, $J_{4',5'}$ = 1.3 Hz, 1H, H-4'), 5.47-5.43 (dd, 1H, $J_{2',3'}$ = 10.7 Hz, $J_{2',1'}$ = 3.4 Hz, 1H, H-2'), 5.41-5.37 (dd, $J_{3',2'}$ = 10.8 Hz, $J_{3',4'}$ =

3.5 Hz, 1H, H-3'), 4.48-4.44 (ddd, $J_{5',4'}$ = 1.1 Hz, $J_{5',6'a}$ = 6.7 Hz, $J_{5',6'b}$ = 7.0 Hz, 1H, H-5'), 4.21-4.16 (A part of an ABX system, $J_{A,B}$ = 11.2 Hz, $J_{A,X}$ = 6.8 Hz, 1H, H-6'a), 4.13-4.08 (B part of an ABX system, $J_{B,A}$ = 11.2 Hz, $J_{A,X}$ = 7.1 Hz, 1H, H-6'b), 2.21-2.02 (m, 12H, 4$COCH_3$)

Synthesis of compound 46

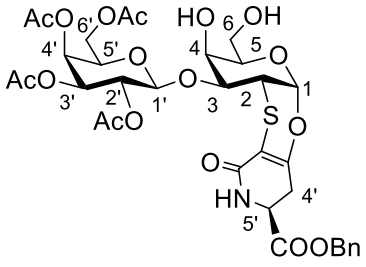


A solution of glycosyl donor **45** (442 mg, 0.90 mmol) and acceptor **25** (200 mg, 0.35 mmol) in dry CH_2Cl_2 (5 mL) was cooled to -30 °C and catalytical TMSOTf (20 µL, 0.11 mmol) was added dropwise. The mixture was stirred for 3 h at room temperature. After completion, 20 µL of triethylamine were added, the solvent was removed under vacuum and the crude was purified by flash chromatography (EtOAc/CH_2Cl_2 6:4). Glycosylation under Schmidt conditions gave high yield of **46** as single β- diastereoisomer (282 mg 90%).

$[\alpha]^{23}_D$ = +78.4 (c 0.25, $CHCl_3$); **ESI-MS *m/z* (%):** 916.25 (100) $[M+Na]^+$, 1808.67 (78) $[2M+Na]^+$; **m.p.** 124-125°C; **^{1}H NMR** (500 MHz, $CDCl_3$) δ: 7.40-7.30 (m, 5H, Bn), 6.05 (s, 1H, NH), 5.58 (d, $J_{1,2}$ = 2.6 Hz, 1H, H-1), 5.30 (d, $J_{4'-3'}$ = 3.5 Hz, 1H, H-4'), 5.25-5.22 (dd, $J_{2'-1'}$ = 7.8 Hz, $J_{2'-3'}$ = 10.4 Hz, 1H, H-2'), 5.28-5.22 (A part of an ABX system, $J_{A,B}$ = 12.3 Hz, 1H, CH_2Ph), 5.21-5.15 (B part of an ABX system, $J_{B,A}$ = 12.0 Hz, 1H, CH_2Ph), 4.99-4.94 (dd, $J_{3'-4'}$ = 3.6 Hz, $J_{3'-2'}$ = 10.5 Hz, 1H, H-3'), 4.68 (m, 1H, H-4), 4.65 (d, $J_{1'-2'}$ = 7.8 Hz, 1H, H-1'), 4.29-4.24 (X part of an ABX system, $J_{X,A}$ = 7.2 Hz, $J_{X,B}$ = 6.8 Hz, 1H, H-5''), 5.26-5.23 (A part of an ABX system, $J_{A,B}$ = 8.1 Hz, 1H, H-6'a), 5.23-5-

20 (B part of an ABX system, $J_{B,A}$ = 8.1 Hz, 1H, H-6'b), 5.2-5.16 (A part of an ABX system, $J_{A,B}$ = 11.4 Hz, 1H, H-6a), 5.16-5.12 (B part of an ABX system, $J_{B,A}$ =12.1 Hz, 1H, H-6b), 3.87 (bs, 1H, H-5), 3.85 (bs, 1H, H-5'), 3.68-3.62 (dd, $J_{2,1}$ = 2.7 Hz, $J_{2,3}$ = 10.7 Hz, 1H, H-2), 3.52-3.47 (dd, $J_{3,4}$ = 2.5 Hz, $J_{3,2}$ = 11.0 Hz, 1H, H-3), 2.87-2.28 (A part of an ABX system, $J_{A,B}$ = 16.0 Hz, $J_{A,X}$ = 7.0 Hz, 1H, H-4''a), 2.67-2.69 (B part of an ABX system, $J_{B,A}$ = 16.7 Hz, $J_{B,X}$ = 7.3 Hz, 1H, H-4''b), 2.18-1.19 (m, 12H, 4 $COCH_3$), 1.07 [s, 9H, $(CH_3)_3C$-Si], 1.05 [s, 9H, $(CH_3)_3C$-Si]; **^{13}C NMR** (125 MHz, $CDCl_3$) δ: 170.0-169.0 (4C_q, $COCH_3$), 169.2 (CO), 164.7 (CO), 154.3 (C_q), 134.6 (C_q, Ph), 128.8 (CH, Ph), 128.8 (2CH, Ph), 128.5 (2CH, Ph),103.7 (C-1'), 97.9 (C_q), 96.4 (C-1), 77.0-66.8 (H-2', H-3', H-4', H-5'), 72.2 (C-4), 69.7 (C-5), 67.1 (C-6'), 67.9 (CH_2Ph), 66.8 (C-6), 66.3 (C-3), 51.5 (C-5''), 39.7 (C-2), 30.9 (C-4''), 27.7 ($C(CH_3)_3$), 27.4 (4CH_3, $COCH_3$), 27.3 ($C(CH_3)_3$), 23.8 (C_q, $(CH_3)_3C$-Si), 20.8 (C_q, $(CH_3)_3C$-Si).

Synthesis of compound 47

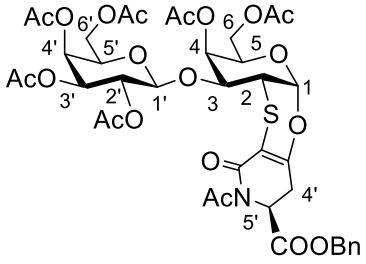


Compound **46** (280 mg, 0.313 mmol) was dissolved in a solution of TBA HF (1M) in dry THF and the solution was stirred overnight at room temperature. After completion, the reaction mixture was concentrated under vacuum to give crude **47** that was used without further purification.

ESI-MS *m/z* (%): 752.17 (100) $[M-H]^-$; **^{1}H NMR** (500 MHz, $CDCl_3$) δ: 7.40-7.30 (m, 5H, 5CH, Ph), 5.89 (s, 1H, NH), 5.61 (d, $J_{1,2}$ = 2.6 Hz, 1H, H-1), 5.40 (d, $J_{4'-3'}$ = 3.5 Hz, 1H, H-4'), 5.25-5.22 (dd, $J_{2'-1'}$ = 7.8 Hz, $J_{2'-3'}$ = 10.4 Hz, 1H, H-2'), 5.28-5.22 (A part of an ABX system, $J_{A,B}$ = 12.3 Hz, 1H, CH_2Ph), 5.21-

5.15 (A part of an ABX system, $J_{B,A}$ = 12.0 Hz, 1H, CH_2Ph), 5.06-5.02 (dd, $J_{3'-4'}$ = 3.4 Hz, $J_{3'-2'}$ = 10.6 Hz, 1H, H-3'), 4.73 (m, 1H, H-4), 4.65 (d, $J_{1'-2'}$ = 8.0 Hz, 1H, H-1'), 4.29-4.24 (X part of an ABX system, $J_{X,A}$ = 7.2 Hz, $J_{X,B}$ = 6.8 Hz, 1H, H-5''), 5.26-5.23 (A part of an ABX system, $J_{A,B}$ = 8.1 Hz, 1H, H-6'a), 5.23-5-20 (B part of an ABX system, $J_{B,A}$ = 8.1 Hz, 1H, H-6'b), 5.2-5.16 (A part of an ABX system, $J_{A,B}$ = 11.4 Hz, 1H, H-6a), 5.16-5.12 (B part of an ABX system, $J_{B,A}$ =12.1 Hz, 1H, H-6b), 3.87 (bs, 1H, H-5), 3.85 (bs, 1H, H-5'), 3.68-3.65 (dd, $J_{2,1}$ = 2.7 Hz, $J_{2,3}$ = 10.7 Hz, 1H, H-2), 3.64-3.58 (dd, $J_{3,4}$ = 2.5 Hz, $J_{3,2}$ = 11.0 Hz, 1H, H-3), 2.58 (s, 1H, OH), 2.87-2.28 (A part of an ABX system, $J_{A,B}$ = 16.0 Hz, $J_{A,X}$ = 7.0 Hz, 1H, H-4''a), 2.67-2.69 (B part of an ABX system, $J_{B,A}$ = 16.7 Hz, $J_{B,X}$ = 7.3 Hz, 1H, H-4''b), 2.58 (s, 1H, OH), 2.22-1.99 (m, 12H, $4CH_3$), 1.55 (s, 1H, OH).

Synthesis of compound 48



Compound **47** (236 mg, 0.31 mmol) was peracetylated in dry CH_2Cl_2 (3 mL) with dry pyridine (780 µL, 9.7 mmol), acetic anhydride (316 µL, 3.32 mmol) and DMAP (2 mg, 0.016 mmol). After stirring for 6 h at room temperature, the mixture was treated with water (1 mL) and washed with H_2O (x1), HCl 1M (x2) and $NaHCO_3$ s.s. (x2). The organic layer was dried over Na_2SO_4 and concentrated under vacuum. The crude was purified by flash chromatography (CH_2Cl_2/CH_3OH 95:5) to give pure peracetylated derivative **48** (198 mg, 73% over two steps).

ESI-MS *m/z* (%): 902.08 (100) $[M+Na]^+$, 918.08 (55) $[M+K]^+$; **^{1}H NMR** (500 MHz, $CDCl_3$) δ: 7.40-7.30 (m, 5H, 5CH, Ph), 5.61 (d, $J_{1,2}$ = 2.6 Hz, 1H, H-1),

5.40 (d, $J_{4'-3'}$ = 3.5 Hz, 1H, H-4'), 5.25-5.22 (dd, $J_{2'-1'}$ = 7.8 Hz, $J_{2'-3'}$ = 10.4 Hz, 1H, H-2'), 5.28-5.22 (A part of an ABX system, $J_{A,B}$ = 12.3 Hz, 1H, CH_2Ph), 5.26-5.23 (A part of an ABX system, $J_{A,B}$ = 8.1 Hz, 1H, H-6'a), 5.23-5-20 (B part of an ABX system, $J_{B,A}$ = 8.1 Hz, 1H, H-6'b), 5.2-5.16 (A part of an ABX system, $J_{A,B}$ = 11.4 Hz, 1H, H-6a), 5.21-5.15 (B part of an ABX system, $J_{B,A}$ = 12.0 Hz, 1H, CH_2Ph), 5.16-5.12 (B part of an ABX system, $J_{B,A}$ = 12.1 Hz, 1H, H-6b), 5.06-5.02 (dd, $J_{3',4'}$ = 3.4 Hz, $J_{3',2'}$ = 10.6 Hz, 1H, H-3'), 4.73 (m, 1H, H-4), 4.65 (d, $J_{1',2'}$ = 8.0 Hz, 1H, H-1'), 4.29-4.24 (X part of an ABX system, $J_{X,A}$ = 7.2 Hz, $J_{X,B}$ = 6.8 Hz, 1H, H-5''), 3.87 (bs, 1H, H-5), 3.85 (bs, 1H, H-5'), 3.68-3.65 (dd, $J_{2,1}$ = 2.7 Hz, $J_{2,3}$ = 10.7 Hz, 1H, H-2), 3.64-3.58 (dd, $J_{3,4}$ = 2.5 Hz, $J_{3,2}$ = 11.0 Hz, 1H, H-3), 2.87-2.28 (A part of an ABX system, $J_{A,B}$ = 16.0 Hz, $J_{A,X}$ = 7.0 Hz, 1H, H-4''a), 2.67-2.69 (B part of an ABX system, $J_{B,A}$ = 16.7 Hz, $J_{B,X}$ = 7.3 Hz, 1H, H-4''b), 2.64-1.96 (m, 21H, $7COCH_3$). **^{13}C NMR** (125 MHz, $CDCl_3$): 171.0 (CO), 171.4 (CO), 170.0-169.0 ($6C_q$, $COCH_3$), 169.6 (CO), 154.3 (C_q), 134.6 (C_q, Ph), 128.8 (CH, Ph), 128.8 (2CH, Ph), 128.5 (2CH, Ph),103.7 (C-1'), 97.9 (C_q), 96.4 (C-1), 77.0-66.8 (H-2', H-3', H-4', H-5'), 72.2 (C-4), 69.7 (C-5), 67.1 (C-6'), 67.9 (CH_2Ph), 66.8 (C-6), 66.3 (C-3), 51.5 (C-5''), 39.7 (C-2), 30.9 (C-4''), 24.8 (CH_3, $NCOCH_3$), 27.4 ($6CH_3$, CH_3).

Part IV

An STnThr mimetic

CHAPTER

10

Synthesis of a novel STnThr mimetic

10.1 Introduction

Sialic acids belong to the family of naturally occurring non-2-ulosonic acids (NulOs),[182–184] that derives from the nine-carbon sugar neuraminic acid (5-amino-3,5-dideoxy-D-glycero-D-galactononulsonic acid).[185] Sialic acids play essential roles in many biological processes and in tumorigenesis through the binding with specific lectins (selectins, siglecs, sialidase, *etc.*).[186–188] N-acetylneuraminic acid (Neu5Ac), a 3-deoxynon-2-ulosonic acid bearing an acetamido group at the C-5 position, is the predominant sialic acid present in mammalian cells that occurs at the end of sugar chains, attached through α2-6 and α2-3 linkages to the surfaces of cells and soluble proteins.[185,189]

The truncated O-glycan, sialyl-Tn antigen (**Chapter 3 - Section 3.2.3**), known as STn, containing a Neu5Ac α-2,6 linked to the Tn antigen backbone, has been shown to be expressed in around 80% of human carcinomas and linked to poor prognosis.[63] Therefore, STn antigen

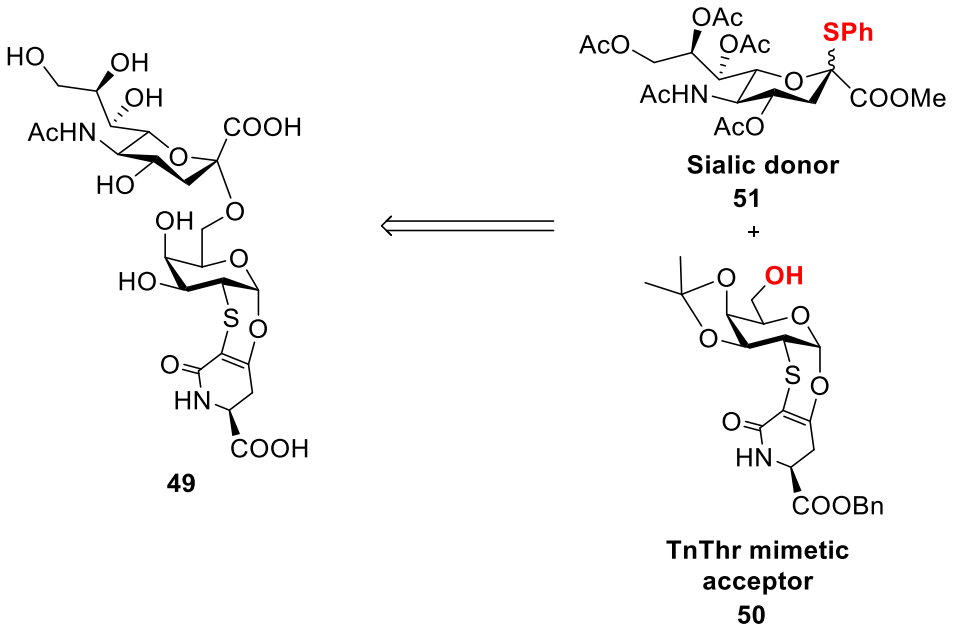
mimetics represents a potential tool to develop cancer vaccines and address the low immunogenicity and instability featuring the natural epitope.

Despite recent advances, chemical sialylation is still considered one of the most difficult glycosylation reactions and a significant synthetic effort is required for the achievement of highly efficient α sialylation.[185,189] Main problems in the synthesis of Neu5Ac derivatives are: a) a low reactivity to glycosylation due to the presence of a strong electron withdrawing group (EWG) at the anomeric center, b) lack of anchimeric assistance by neighboring groups at C- 3, c) lack of hydrogens at C-2 which makes challenging the determination of the anomeric configuration

To take advantage from the advantages of the TnThr mimetic[82] structure (**Chapter 5**) and to synthesize the third mucin-related antigen mimetic, the synthesis of a physiological stable STnThr mimetic was carefully planned. Moreover, to demonstrate that **49** behaves as a real mimetic of the natural STnThr, (*i.e.*, it is presented to and recognized by its natural targets), and keeping in mind the debate on the specificity of monoclonal antibodies *vs.* glycopeptide targets,[190,191] the STnThr mimetic was screened as a substrate for the N1 type neuraminidase of the *Influenza* virus A. Specifically, the effect of the structurally constrained galactoside moiety on the substrate specificity of neuraminidase was investigated.[192]

10.2 Results and Discussion

10.2.1 Synthesis of an STnThr mimetic

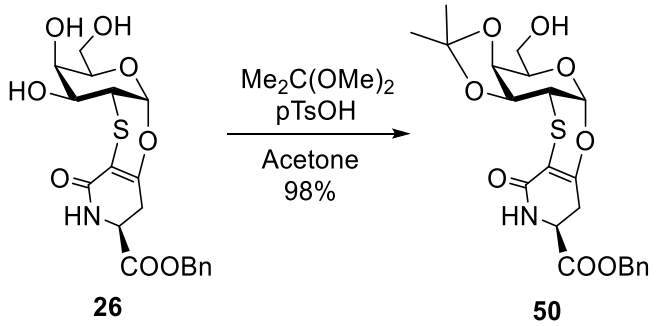


Scheme 10.2.1.1 Retrosynthesis of the STnThr mimetic **49**

For the preparation of **49**, a synthetic pathway was devised (**Scheme 10.2.1.1**) starting from a protected derivative of the TnThr mimetic, bearing free primary 6-hydroxyl group (**50**), and a fully protected Neu5Ac donor properly activated at the anomeric center (**51**).

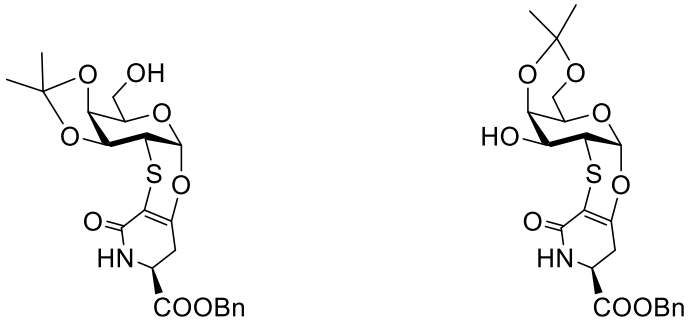
An isopropylidene ketal was chosen as protective group for the TnThr mimetic acceptor (**50**). Unprotected **26** was reacted in the presence of 2,2-dimethoxypropane and a catalytical amount of *p*-toluenesulfonic acid (*p*TsOH) in acetone until complete conversion into the 3,4 acetonide **50** (**Scheme 10.2.1.2**). The progress of the reaction was monitored by NMR, highlighting the initial formation of the kinetically favored 4,6 isomer **52**

followed by the conversion into the thermodynamic, more stable, 3,4 five members dioxolane (**Scheme 10.2.1.3**). Compound **50** was isolated after flash chromatography purification in a 98% yield.



26 **50**

Scheme 10.2.1.2 Selective protection of the TnThr mimetic **26**



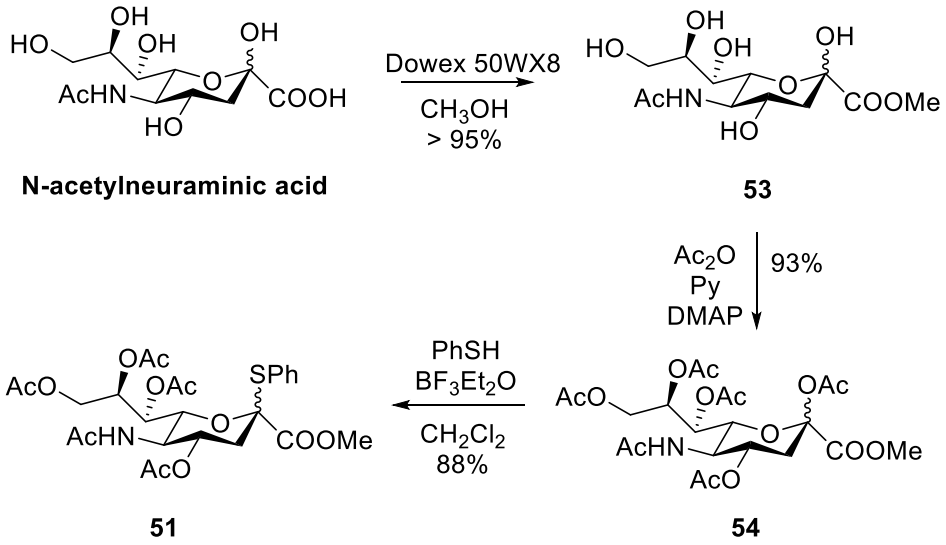
50
Thermodynamic product

52
Kinetic product

Scheme 10.2.1.3 Thermodynamic (**50**) and kinetic (**52**) products

Regarding the synthesis of Neu5Ac acceptor, the initial idea concerned the use of a methyl ester protecting group for the carboxylic acid (**51 - Scheme 10.2.1.4**). N-acetylneuraminic acid was efficiently converted into the methylated derivative through an acid catalyzed Fisher esterification with Dowex® 50WX8 in MeOH affording **53** in a > 95% yield as predominant β anomer (> 90%). Subsequently, **53** was acetylated under standard

conditions[193] and after flash chromatography purification the peracetylated derivative **54** was converted, as reported by Marra, Sinaÿ *et al*,[194] in the presence of thiophenol and boron trifluoride etherate ($BF_3 \cdot Et_2O$) into the thio-donor **51** in a 88% yield.

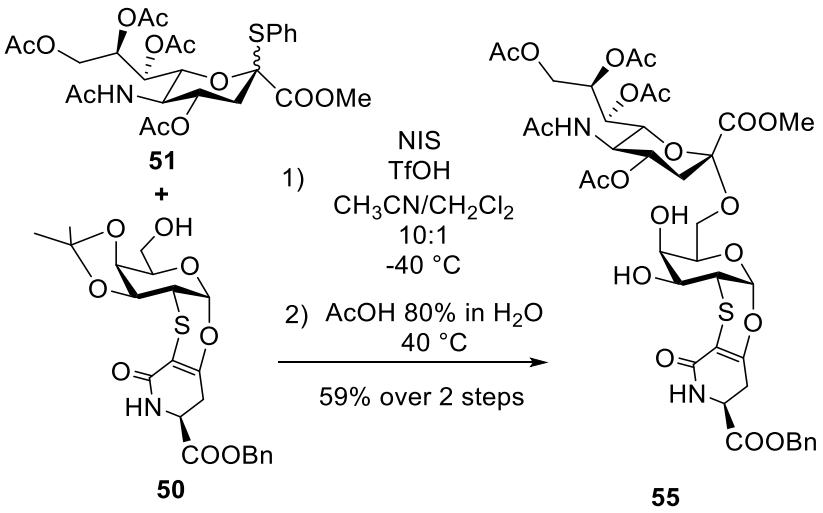


N-acetylneuraminic acid **53** **54** **51**

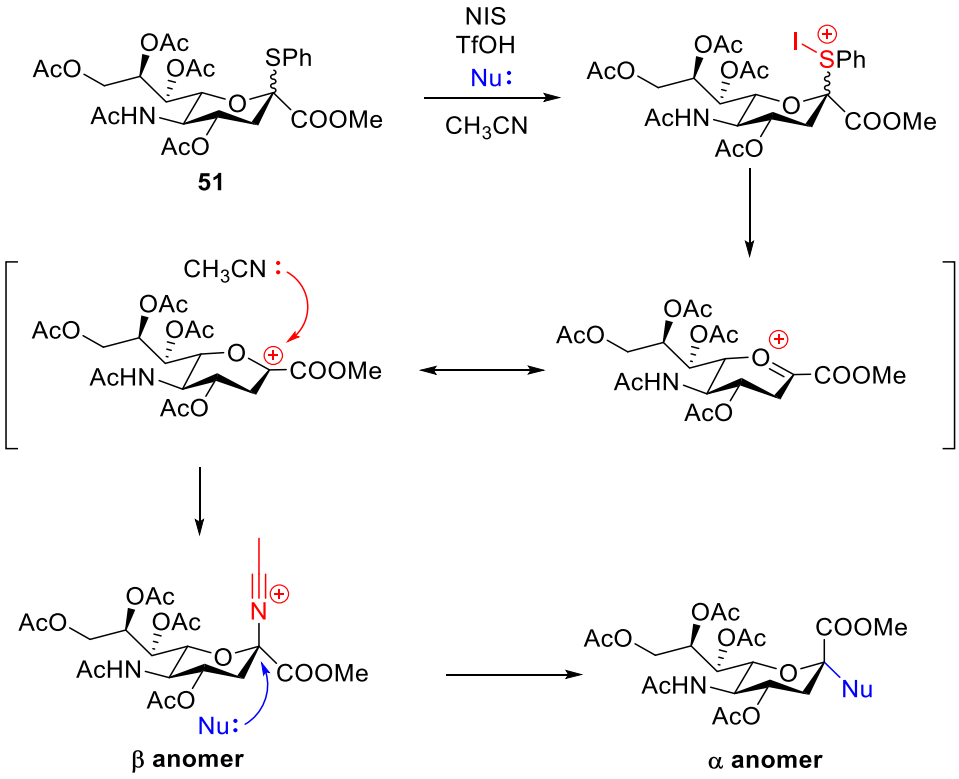
Scheme 10.2.1.4 Synthesis of the methyl ester donor **51**

The glycosyl donor **51** was thus reacted with **50** in the presence of N-iodosuccinimide (NIS) and trifluoromethanesulfonic acid (TfOH).

A first attempt of glycosylation was performed in CH_2Cl_2 affording **55** in a very low stereoselectivity (α/β 6:4). Switching to the participating CH_3CN solvent, at low temperature (-40 °C), allowed to increase the selectivity and the α anomer up to 90%. After isopropylidene cleavage with an 80% solution of AcOH in H_2O and flash chromatography purification, pure α isomer **55** was isolated in a 59% yield over 2 steps (**Scheme 10.2.1.5**).



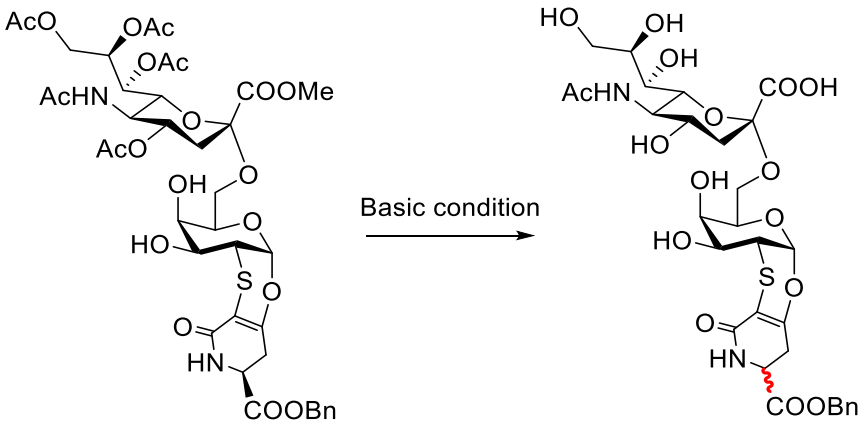
Scheme 10.2.1.5 Synthesis of glycoside **55**



Scheme 10.2.1.6 Mechanism of activation of **51** with NIS in CH_3CN as solvent.

Indeed, in the activation mechanism of thio-Neu5Ac donors by NIS and TfOH[195], an electrophilic I^+ generated *in situ* reacts with the lone pair of the anomeric sulfur of **51**. The sulfonium moiety, an excellent leaving group, can be displaced first by acetonitrile to give a nitrilium ion that adopts a preferred axial (β) configuration and the following nucleophilic substitution of the nitrilium ion by the hydroxyl group of the glycosidic acceptor gives predominantly equatorial (α)-glycosides (**Scheme 10.2.1.6**).

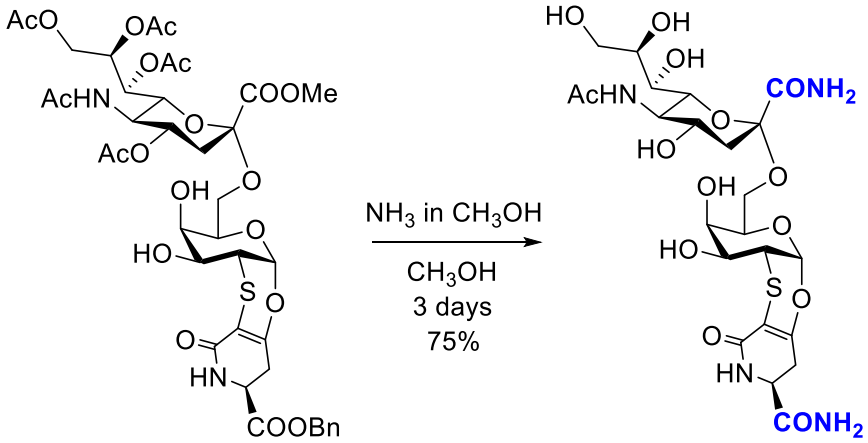
Then, deacetylation of **55** was initially performed in manifold basic conditions (K_2CO_3 in CH_3OH or H_2O, NaOH in CH_3OH or H_2O, $NaOCH_3$ in CH_3OH) but in all cases, the too strong basic media caused not only deacetylation and deprotection of the methyl ester but also an isomerization at the lactam ring level, as already seen in first attempt to synthesize the TFThr mimetic[82] (**Scheme 10.2.1.7**).



55

Scheme 10.2.1.7 Deacetylation in strong basic condition of **55**

Therefore, a mild deacetylation was conducted upon treatment with a solution of NH_3 in CH_3OH (2M) (**Scheme 10.2.1.8**). Cleavage of all the acetyl groups needed prolonged reaction time which unfortunately resulted in the transformation of both carboxylic groups into amide residues. The diamide derivative **56** was thus isolated after flash chromatography, in a 75% yield.

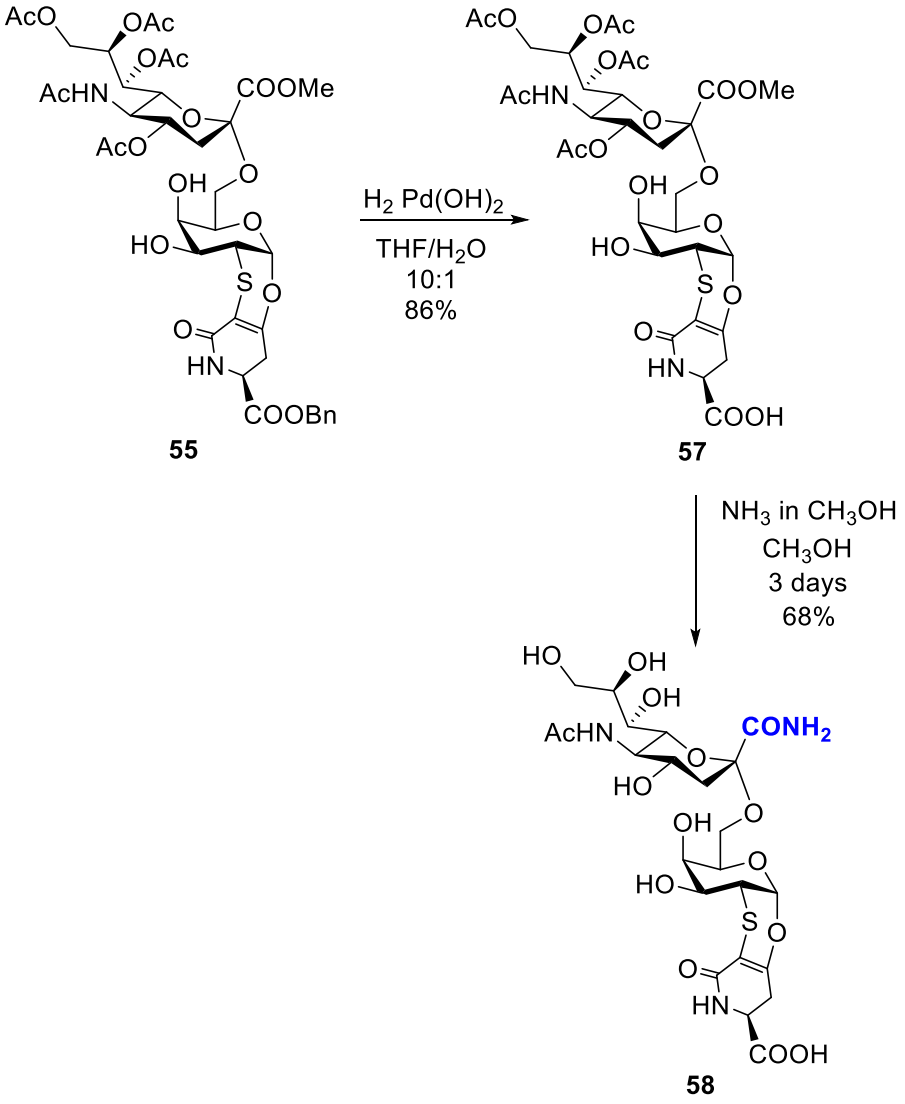


55 **56**

Scheme 10.2.1.8 Deacetylation of **55** and synthesis of **56**

However, since the carboxylic group is essential for physiological activity of sialic derivatives, different strategies were explored. Deacetylation with NH_3 in CH_3OH is efficient but the formation of amides must be avoided. In this regard, **55** was subjected to a palladium catalyzed hydrogenation and isolated, free carboxylic derivative, **57** was subsequent treated with a solution of NH_3 in CH_3OH (2M) (**Scheme 10.2.1.9**). From these steps a single product (**58**) in which only the methyl ester group was converted in amide was isolated in a 68 % yield. This experimental evidence led to the

idea, that a protection with benzyl esters for both carboxylic acids could be the final solution.

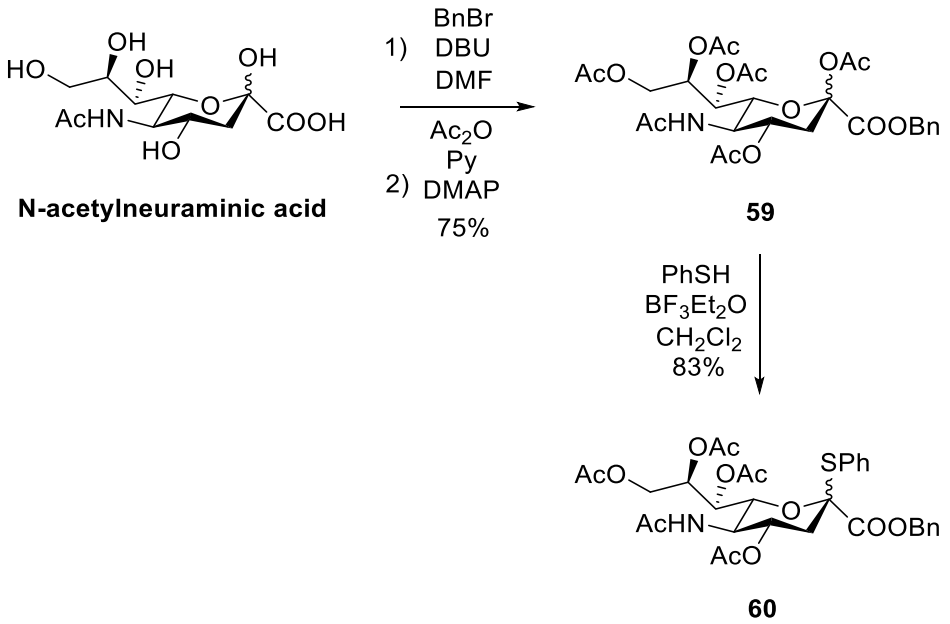


Scheme 10.2.1.9 Palladium catalyzed hydrogenation and deacetylation with NH_3 in CH_3OH of **55**

In light of the above, the synthesis of the Neu5Ac donor was revised (**Scheme 10.2.1.10**). Esterification of Neu5Ac with benzyl bromide (BnBr)

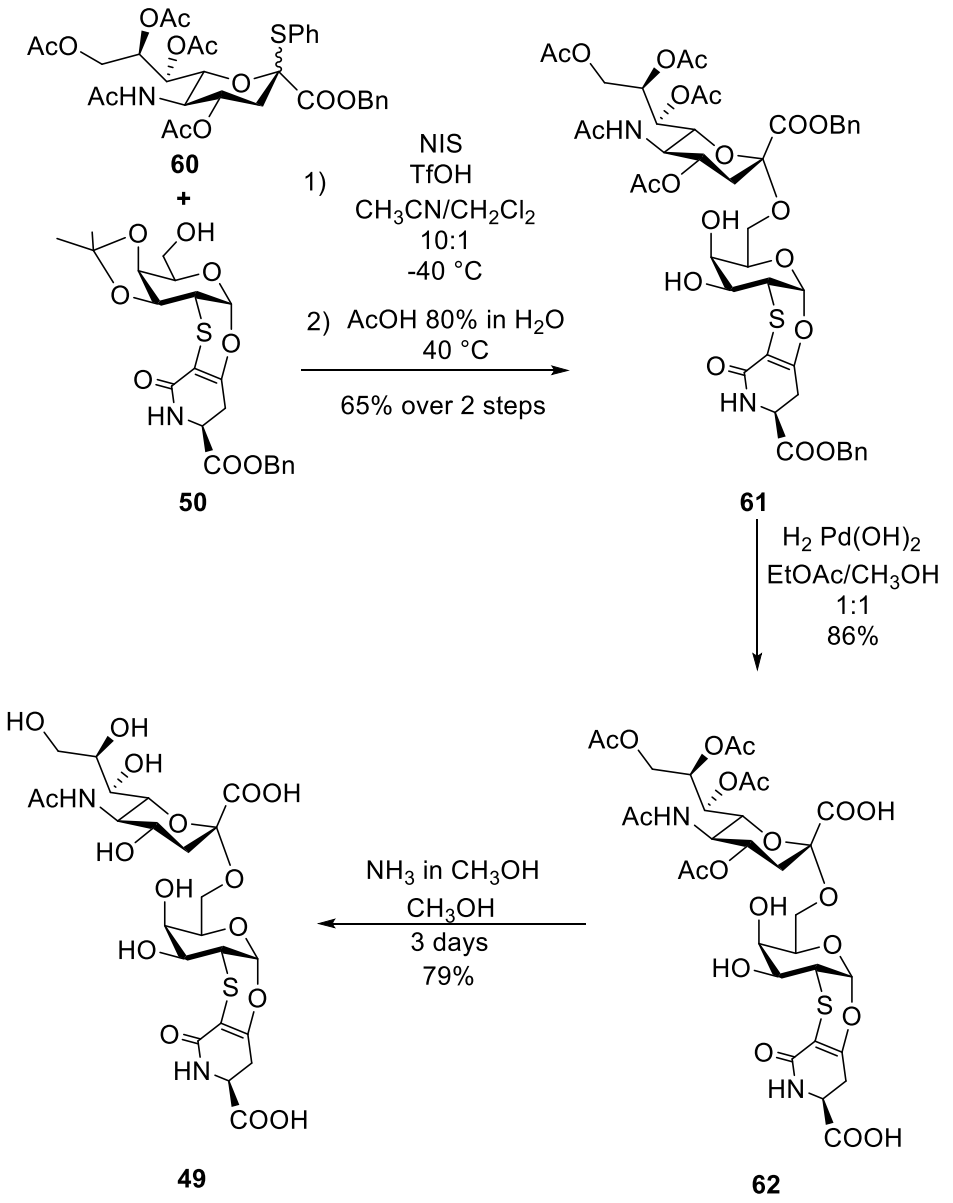
and 1,5-diazabiciclo(5.4.0)undec-7-ene (DBU) in DMF followed by an acetylation under standard condition allowed to isolate the pentaacetylated benzyl ester **59** in a 75 % yield over 2 steps as predominant β anomer (> 90%).

Subsequently, **59** was converted in the corresponding thio-derivative **60** in an 83% yield (β > 90%).



Scheme 10.2.1.10 Synthesis of the benzyl ester donor **60**

Therefore, in the same conditions reported before, from the new glycosyl donor **60** alternative dibenzyl ester STnThr derivative (**61**) was synthesized in a 65 % yield over 2 steps. **61** was debenzylated with $Pd(OH)_2$ in H_2 atmosphere affording intermediate **62**, bearing both free carboxylic acids, in an 86 % yield. Finally, deacetylation with NH_3 in CH_3OH (2M) of **62** led to the synthesis of the fully unprotected STnThr mimetic **49**, as a white solid, in a 79% yield. (**Scheme 10.2.1.11**).



Scheme 10.2.1.11 Final synthesis of the STnThr mimetic **49**

10.2.2 Recognition of the STnThr mimetic by *Influenza* neuraminidase

In the search of serological biomarkers for clinical diagnosis of cancer, the substrate specificity of neuraminidase is used to identify sialylated antigens.[192] The widely studied *Influenza* neuraminidase cleaves α2-3- and α2-6 sialic acids. The enzyme acts on the α-anomer of the sialic acid linked to the proximal sugar and releases the free sialic acid.[196] Thus compound **49** was screened as a potential Influenza neuraminidase inhibitor to investigate the effect of the structurally constrained galactoside moiety on the substrate specificity.

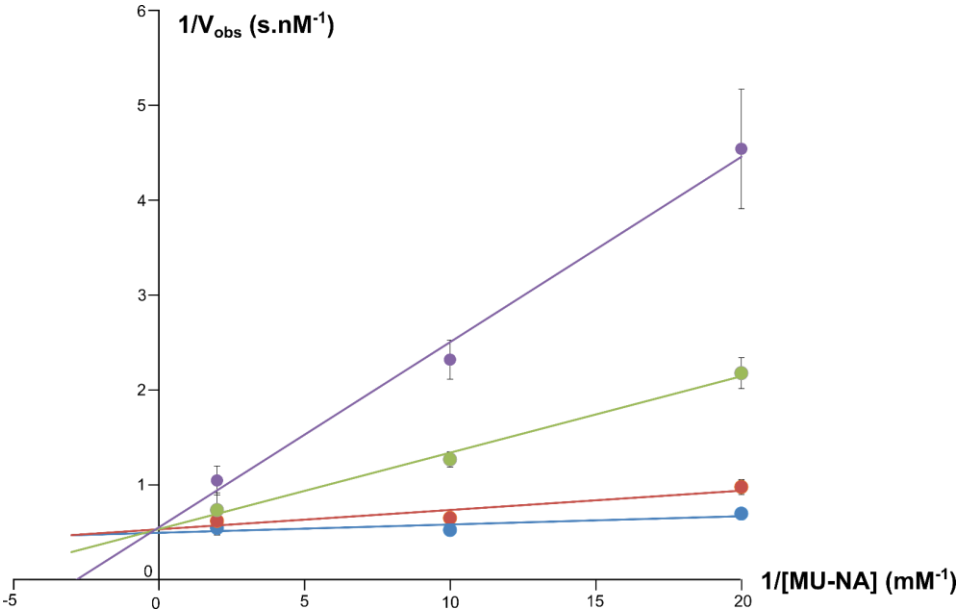


Figure 10.2.2.1 Lineweaver Burke plot of *Influenza* A neuraminidase N1 by compound **49**. Data were obtained using a concentration of inhibitor of 0 µM (blue), 100 µM (red), 500 µM (green) and 1000 µM (purple). Data are depicted as Mean ± Standard deviation from three independent experiments.

Using recombinant Neuraminidase (N1) from H1N1 strain (A/Brevig Mission/1/1918(H1N1)) **49** was indeed found to be a competitive inhibitor, as appreciated from the inhibition plot (**Figure 10.2.2.1**).

The inhibition constant (*Ki)* was calculated to be 40 ± 10 µM, a value markedly low compared to the sub-nanomolar inhibition constants found for drugs oseltamivir or zanamivir,[197] yet it indicates that **49** can effectively bind to the (N1) active site. Indeed, standard fluorogenic substrate for influenza neuraminidase, 2'-(4-methylumbelliferyl)-alpha-D-N-acetylneuraminic acid (MU-NA) shows a comparable Michaelis–Menten constants (K_M) of 20 to 60 µM[198,199] (80 µM calculated in our assay conditions).

Using the X-ray structure of N1 bound to zanamivir (PDB 3B7E)[200] as a template, a model of N1-**49** complex was built (**Figure 10.2.2.2**) and analyzed by a 5 ns molecular dynamics simulations. A set of stabilizing interactions occurs between the sialic acid moiety of **49** with N1 active site residues (**Figure 10.2.2.2**). Arg 293 and Arg 368 interact with the negatively charged located on the C-1 carboxylate group, whereas two other arginines, Arg 152 and Arg 225 are hydrogen bonded to the C-8 atoms. Glu 119 completes this binding, as it interacts through a hydrogen bond with the C-4 atom. In contrast, the aglycon of **49** does not interact with the protein residues and points out toward the solvent. The observed lack of interactions of the aglycon of **49** can explain the micromolar range of the *in vitro* affinity constant, which is solely due to the binding of sialic acid to the active site. The described results demonstrate that the TnThr mimetic portion of **49**, does not prevent recognition of the sialic acid

moiety by the investigated carbohydrate-binding protein, suggesting potential recognition by other hydrolytic enzymes or lectins.

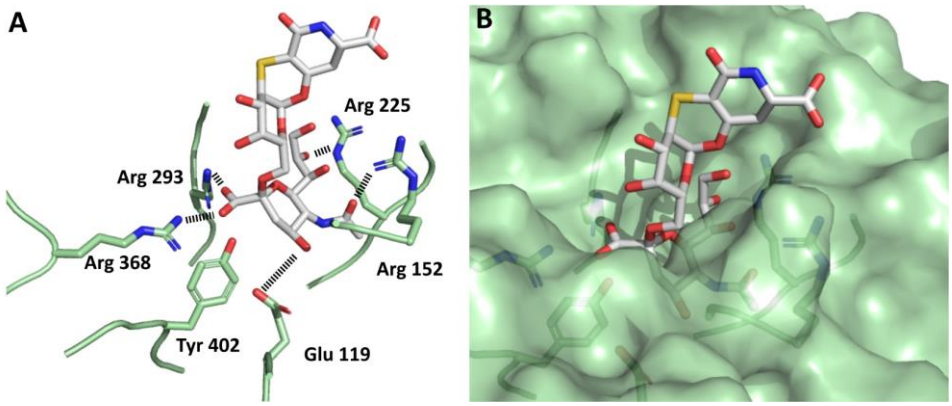


Figure 10.2.2.2 (**A**) Binding interactions between compound **49** and Neuraminidase N1 active site residues. (**B**) Active site pocket of neuraminidase N1 complexed with compound **49**.

10.3 Conclusions

In conclusion, in this chapter a simple, highly stereoselective and effective strategy for the development of an STnThr mimetic was reported. **49** was isolated as single α anomer, in overall yield, considering glycosylation and deprotections reactions, of 44% over 4 steps.

As for native STn and sialyl oligosaccharides, the STnThr mimetic is recognized by neuraminidase. In particular, a model of neuraminidase N1 from H1N1 complexed with **49** showed the sialic moiety nested in the N1 active site, proving that the rigid aglycone did not affect the presentation of the sialyl portion to the enzyme.

Hence, we planned to employ this new mimetic, as monovalent construct and in multivalent architectures, to screen the interaction with natural targets (selectins, siglecs, sialidase, *etc.*).[186–188]

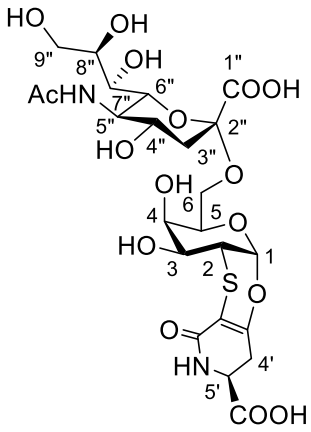
10.4 Experimental section

10.4.1 General methods

Analytical grade solvents and commercially available reagents were used without further purification. For anhydrous reactions, solvent stored over 3 Å molecular sieves were used. Silica gel flash column chromatography purifications were performed using Geduran® Si 60 (0.040-0.063 mm). TLC analyses were performed on glass Merck silica gel 60 F_{254} plates. ^{1}H NMR, ^{13}C NMR and 2D-NMR spectra were recorded on a 500 MHz Bruker AVANCE II at 298 K, unless otherwise stated. All chemical shifts are reported in parts per million (δ) referenced to residual nondeuterated solvent. Multiplicity abbreviation: b = broad, s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet were used. ESI-MS spectra were carried out on a linear ion-trap double quadrupole mass spectrometer using electrospray ionization (ESI) technique (LTQ-XL - Thermo Fisher) . Optical rotation measurements were performed on a JASCO DIP-370 polarimeter. Melting point were recorded on a BUCHI 510.

10.4.2 Synthesis

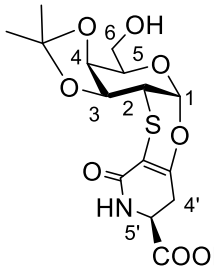
Synthesis of compound 49



Compound **62** (43 mg, 54.24 µmol) was solubilized in NH_3 in CH_3OH 4M (1 mL) and the reaction was stirred at room temperature. After 4 days, the solvent was removed and the product was purified by several washing with Et_2O to give pure **49** (27 mg, 79%).

ESI-MS *m/z* (%): 623.42 (100) $[M]^-$, 311.25 (98) $[M]^{2-}$, **$[\alpha]^{25}_D$ =** +109.5 (c 0.001, H_2O), **^{1}H NMR** (500 MHz, D_2O) δ: 5.72 (d, 1H, $J_{1,2}$= 2.7 Hz, H-1), 4.23-4.18 (m, 1H, H-5), 4.07-4.03 (m, 1H, H-5'), 3.99-3.95 (m, 1H, H-4), 3.92-3.86 (m, 1H, H-6_a), 3.86-3.71 (m, 4H, H-3, H-5'', H-8'', H-9''$_a$), 3.68-3.48 (m, 5H, H-4'', H-6_b, H-6'', H-7'', H-9''$_b$), 3.41 (dd, 1H, $J_{2,1}$= 2.7 Hz, $J_{2,3}$= 11.2 Hz, H-2), 2.89 (dd, 1H, $J_{4'a,4'b}$= 16.9 Hz, $J_{4'a,5'}$= 6.8 Hz, H-4'$_a$), 2.72-2.58 (m, 2H, H-4'$_b$,H-3''$_{eq}$), 1.95 (s, 3H, CH_3), 1.67-1.59 (m, 1H, H-3''$_{ax}$), **^{13}C NMR** (125 MHz, D_2O) δ: 177.3 (C_q), 174.9 (C_q), 173.4 (C_q), 167.5 (C_q), 158.3 (C_q), 100.4 (C_q), 96.0 (C-1), 93.7 (C_q), 72.6 (C-6''), 71.9 (C-5), 71.7 (C-8''), 68.4 (C-4), 68.2 (C-7'', C-4''), 65.0 (C-3), 63.7 (C-6), 63.0 (C-9''), 52.8 (C-5'), 51.8 (C-5''), 40.1 (C-3''), 38.0 (C-2), 31.1 (C-4'), 22.0 (CH_3)

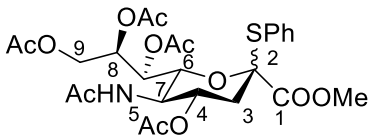
Synthesis of compound 50



Compound **26** (600 mg, 1.42 mmol) was suspended in acetone (15 mL), $Me_2C(OMe)_2$ (10.3 g, 99 mmol) and *p*TsOH (pH 3) were added and the solution was stirred at room temperature for 5 h. After complete conversion, the mixture was quenched with Et_3N (pH 8), diluted with CH_2Cl_2 and washed with HCl 1M (x4). The organic layer was dried over Na_2SO_4 and concentrated under vacuum. The crude was purified by flash chromatography (EtOAc/CH_3OH 9:1) to give pure **50** as white solid (628 mg, > 95%).

ESI-MS *m/z* (%): 486.17 (100) $[M+Na]^+$, 502.17 (90) $[M+K]^+$, **$[\alpha]^{26}_D$ =** +118.5 (c 0.001, CH_3OD), **^{1}H NMR** (500 MHz, $DMSOd_6$) δ: 7.80 (d, 1H, $J_{NH,5'}$ 3.2 Hz, NH), 7.42-7.31 (m, 5H, Bn), 5.48 (d, 1H, $J_{1,2}$= 2.9 Hz, H-1), 5.22-5.16 (m, 2H, CH_2Bn), 4.38-4.32 (m, 1H, H-5'), 4.24-4.19 (m, 1H, H-5), 4.18-4.15 (m, 1H, H-4), 4.07 (dd, 1H, $J_{3,2}$= 8.7 Hz, $J_{3,4}$= 5.0 Hz, H-3), 3.66-3.55 (m, 2H, H-6_a, H-6_b), 3.17 (dd, 1H, $J_{2,3}$= 8.7 Hz, $J_{2,1}$= 2.9 Hz, H-2), 3.06 (dd, 1H, $J_{4'a,4'b}$= 16.7 Hz, $J_{4'a,5'}$= 7.1 Hz, H-4'$_a$), 2.61 (dd, 1H, $J_{4'b,4'a}$= 16.7 Hz, $J_{4'b,5'}$= 5.0 Hz, H-4'$_b$), 1.45 (s, 3H, CH_3), 1.27 (s, 3H, CH_3). **^{13}C NMR** (125 MHz, $DMSOd_6$) δ: 171.7 (C_q), 154.7 (C_q), 155.5 (C_q), 136.2 (C_q), 128.9 (CH, Bn), 128.6 (CH, Bn), 128.2 (CH, Bn), 109.1 (C_q), 96.6 (C_q), 95.2 (C-1), 72.5 (C-3), 72.4 (C-4), 69.9 (C-5), 66.9 (CH_2Ph), 60.9 (C-6), 51.0 (C-5'), 39.11 (C-2), 30.78 (C-4'), 28.51 (CH_3), 26.61 (CH_3)

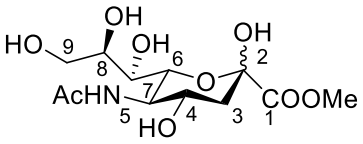
Synthesis of compound 51



To a solution of **54** (3.25 g, 6.09 mmol) in CH_2Cl_2 dry (50 mL), cooled to 0 °C, PhSH (730 mg, 6.63 mmol) and $BF_3{\cdot}Et_2O$ (1.01 g, 7.09 mmol) were added and the solution was stirred at room temperature overnight. After complete conversion, the mixture was diluted with CH_2Cl_2 and washed with $NaHCO_3$ s.s. (x5). The organic layer was dried over Na_2SO_4 and concentrated under vacuum. The crude was purified by flash chromatography (CH_2Cl_2/CH_3OH 9:1) to give **51** as yellow solid (3.10 g, 88%, β anomer > 90%).

^{1}H NMR (β anomer) (500 MHz, $CDCl_3$) δ: 7.46-7.19 (m, 5H, SPh), 5.56 (d, 1H, $J_{NH,5}$= 10.3 Hz NH), 5.46 (dd, 1H, $J_{7,6}$= 2.5 Hz, $J_{7,8}$= 2.3 Hz, H-7), 5.41-5.34 (m, 1H, H-4), 4.94 (ddd, 1H, $J_{8,9b}$= 8.6 Hz, $J_{8,9a}$= 2.3 Hz, $J_{8,7}$= 2.3 Hz, H-8), 4.62 (dd, 1H, $J_{6,5}$= 10.5 Hz, $J_{6,7}$= 2.5 Hz, H-6), 4.48 (dd, 1H, $J_{9a,9b}$= 12.3 Hz, $J_{9a,8}$= 2.3 Hz, H-9a), 4.12 (ddd, 1H, $J_{5,6}$= 10.5 Hz, $J_{5,4}$= 10.4 Hz, $J_{5,NH}$= 10.3 Hz, H-5), 3.99 (dd, 1H, $J_{9b,9a}$= 12.3 Hz, $J_{9b,8}$= 8.6 Hz, H-9b), 3.58 (s, 3H, OCH_3), 2.66 (dd, 1H, $J_{3eq,3ax}$= 13.8 Hz, $J_{3eq,4}$= 4.8 Hz, H-3_{eq}), 2.16-2.11 (m, 1H, H-3_{ax}), 2.09 (s, 3H, CH_3), 2.07 (s, 3H, CH_3), 2.03 (s, 3H, CH_3), 1.95 (s, 3H, CH_3), 1.89 (s, 3H, CH_3), **^{13}C NMR (β anomer)** (125 MHz, $CDCl_3$) δ: 171.3 (C_q), 171.1 (C_q), 170.4 (C_q), 170.4 (C_q), 170.3 (C_q), 168.3 (C1), 136.3 (CH, SPh) 129.9 (CH, SPh), 129.2 (CH, SPh), 129.0 (C_q, SPh), 89.1 (C-2), 73.3 (C-8, C-6), 69.2 (C-4), 69.0 (C-7), 62.8 (C-9), 52.7 (OCH_3), 49.5 (C-5), 37.6 (C-3), 23.3 (CH_3), 21.2 (CH_3), 21.0 (CH_3), 20.9 (CH_3), 20.8 (CH_3).

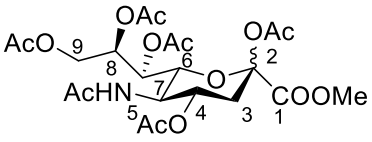
Synthesis of compound 53



N-acetylneuraminic acid (5.05 g, 16.31 mmol) was suspended in CH_3OH (170 mL), Dowex 50WX8 (1.5 g) was added and the suspension was stirred at 40 °C. After complete dissolution, Dowex 50WX8 was filtered off and the organic layer was concentrated, under reduce pressure, to give **53** as white solid (5.06 g, > 95%, β anomer > 90%).

ESI-MS *m/z* (%): 346.17 (100) $[M+Na]^+$, 668.92 (65) $[M+Na]^{2+}$, 386.17 (55) $[M+K]^+$, 708.83 (35) $[M+K]^{2+}$, **^{1}H NMR (β anomer)** (500 MHz, CD_3OD) δ: 4.08-3.99 (m, 1H, H-4), 3.99 (dd, 1H, $J_{6,5}$= 10.1 Hz, $J_{6,7}$= 1.4 Hz, H-6), 3,84-3,76 (m, 5H, H-5, OCH_3, H-9_a), 3.72-3.68 (m, 1H, H-8), 3.62 (dd, 1H, $J_{9b,9a}$= 11.2 Hz, $J_{9b,8}$= 5.7 Hz, H-9_b), 3.48 (dd, 1H, $J_{7,8}$= 9.1 Hz, $J_{7,6}$= 1.4 Hz, H-7), 2.22 (dd, 1H, $J_{3eq,3ax}$= 12.9 Hz, $J_{3eq,4}$= 5.0 Hz, H-3_{eq}), 2.00 (s, 3H, NHAc), 1.89 (dd, 1H, $J_{3ax,3eq}$= 12.9 Hz, $J_{3ax,4}$= 11.6 Hz, H-3_{ax}), **^{13}C NMR (β anomer)** (125 MHz, CD_3OD) δ: 175.1 (C_q - NHAc), 171.8 (C-1), 96.7 (C-2), 72.1 (C-6), 71.7 (C-8), 70.2 (C-4), 67.9 (C-7), 64.9 (C-9), 54.3 (OCH_3), 53.1 (C-5), 40.7 (C-3), 22.6 (NHAc).

Synthesis of compound 54

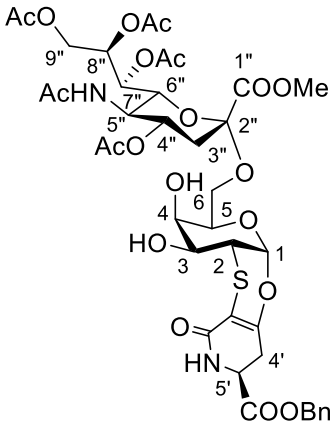


To a solution of **53** (5.05 g, 15.62 mmol) in pyridine (80 mL), Ac_2O (10.3 g, 99 mmol) and DMAP (100 mg) were added and the solution was stirred at room temperature overnight. After complete conversion, the mixture was diluted with CH_2Cl_2 and washed with HCl 1M (x5). The organic layer was dried over Na_2SO_4 and concentrated under

vacuum. The crude was purified by flash chromatography (CH_2Cl_2/ CH_3OH 9:1) to give **54** as white solid (7.73 g, 93%, β anomer > 90%).

ESI-MS *m/z* (%): 556.17 (100) $[M+Na]^+$ **^{1}H NMR (β anomer)** (500 MHz, $CDCl_3$) δ: 5.40-5.35 (m, 2H, NH, H-7), 5.27-5.20 (m, 1H, H-4), 5.07-5.04 (m, 1H, H-8), 4.48 (dd, 1H, $J_{9a,9b}$= 12.4 Hz, $J_{9a,8}$= 2.6 Hz, H-9_a), 4.14-4.07 (m, 3H, H-5, H-6, H-9_b), 3.80 (s, 3H, OCH_3), 2.53 (dd, 1H, $J_{3eq,3ax}$= 13.6 Hz, $J_{3eq,4}$= 5.0 Hz, H-3_{eq}), 2.34 (s, 3H, CH_3), 2.13 (s, 3H, CH_3), 2.11-2.07 (m, 1H, H-3_{ax}), 2.09 (s, 3H, CH_3), 2.02 (s, 3H, CH_3), 2.02 (s, 3H, CH_3), 1.88 (s, 3H, CH_3), **^{13}C NMR (β anomer)** (125 MHz, $CDCl_3$) δ: 171.1 (C_q), 171.1 (C_q), 170.7 (C_q), 170.4 ($2C_q$), 170.3 (C_q), 168.3 (C_q), 166.4 (C_q), 97.6 (C-2), 72.97 (C-6), 71.5 (C-8), 68.4 (C-4), 67.9 (C-7), 62.2 (C-9), 53.3 (OCH_3), 49.4 (C-5), 36.0 (C-3), 23.3 (CH_3), 21.1 (CH_3), 21.0 (CH_3), 20.9 ($2CH_3$), 20.8 (CH_3).

Synthesis of compound 55

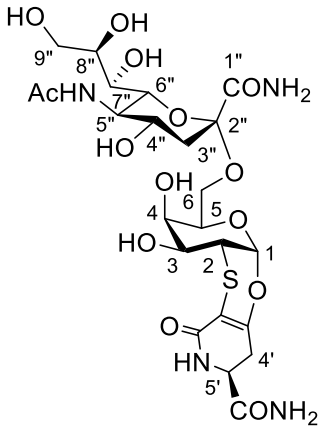


To a suspension of **50** (500 mg, 1.08 mmol) and **51** (1.50 g, 2.57 mmol) in a mixture of CH_3CN/CH_2Cl_2 dry 10:1 (10 mL) cooled to - 40 °C, NIS (1.04 g, 4.62 mmol) and TfOH (203 mg, 1.35 mmol) were added and the solution was stirred at - 40 °C under an N_2 atmosphere. After 2 h, the reaction was quenched with Et_3N (pH 8), diluted with CH_2Cl_2 and washed with $Na_2S_2O_3$ 1 M (x3). The organic layer was dried over Na_2SO_4 and concentrated under vacuum. The crude was solubilized in AcOH 80% in H_2O (20 mL) and the reaction was stirred overnight at 40 °C. The solvent was removed *in vacuo* and the product was purified by flash

chromatography (EtOAc/CH_3OH 98:2) to give pure **55** (571 mg, 59% over 2 steps)

ESI-MS *m/z* (%): 919.42 (100) $[M+Na]^+$, 935.42 (90) $[M+K]^+$, **^{1}H NMR (α anomer)** (500 MHz, $CDCl_3$) δ: 7.37-7.27 (m, 5H, Bn), 6.71 (bs, 1H, NH), 5.71 (d, 1H, $J_{NH,5''}$= 9.8 Hz, NH), 5.57 (d, 1H, $J_{1,2}$= 2.7 Hz, H-1), 5.37-5.33 (m, 1H, H-8''), 5.31-5.28 (m, 1H, H-7''), 5.21-5.14 (m, 2H, CH_2Bn), 4.91-4.81 (m, 1H, H-4''), 4.38- 4.27 (m, 2H, H-5', H-9''$_a$), 4.19-3.98 (m, 5H, H-4, H-5, H-5'', H-6'', H-9''$_b$), 3.95-3.86 (m, 1H, H-6$_a$), 3.78 (s, 1H, OCH_3), 3.72-3.63 (m, 2H, H-3, H-6$_b$), 3.48 (dd, 1H, 1H, $J_{2,1}$= 2.7 Hz, $J_{2,3}$= 10.7 Hz, H-2), 2.95-2.87 (m, 1H, H-4''$_a$), 2.80-2.71 (m, 1H, H-4'$_b$), 2.57 (dd, 1H, $J_{3''eq,3''ax}$= 12.8 Hz, $J_{3''eq,4''}$= 4.6 Hz, H-3''$_{eq}$), 2.10 (s, 3H, CH_3), 2.09 (s, 3H, CH_3), 1.98-1.90 (m, 1H, H-3''$_{ax}$), 2.00 (s, 3H, CH_3), 1.99 (s, 3H, CH_3), 1.85 (s, 3H, CH_3), **^{13}C NMR (α anomer)** (125 MHz, $CDCl_3$) δ: 171.0, 170.6, 170.4, 170.2, 169.8, 168.2, 165.8, 155.9, 134.9, 128.8, 128.4, 98.8, 96.5 (C-1), 96.3, 72.8, 71.3, 69.2 (C-8''), 69.1 (C-4''), 67.9, 67.8 (CH_2Bn), 67.6 (C-7''), 65.7 (C-3), 63.2 (C-6), 62.7 (C-9''), 53.1 (OCH_3), 51.4 (C-5'), 49.2, 39.1 (C-2), 37.4 (C-3''), 30.7 (C-4'), 23.2 (CH_3), 21.1 (CH_3), 20.9 (CH_3), 20.9 (CH_3), 20.8 (CH_3).

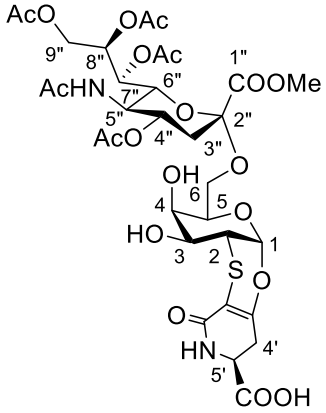
Synthesis of compound 56



Compound **55** (50 mg, 55.75 µmol) was solubilized in NH_3 in CH_3OH 4M (1 mL) and the reaction was stirred at room temperature. After 4 days, the solvent was removed and the product was purified by several washing with Et_2O to give pure **56** (26 mg, 75%).

ESI-MS *m/z* (%): 621.33 (100) $[M]^-$, **$[\alpha]^{25}_D$ =** +111.2 (c 0.001, H_2O), **^{1}H NMR** (500 MHz, D_2O) δ: 5.85 (d, 1H, $J_{1,2}$= 2.7 Hz, H-1), 4.37-4.34 (m, 1H, H-5'), 4.33-4.28 (m, 1H, H-5), 4.10-4.07 (m, 1H, H-4), 4.03-3.94 (m, 2H, H-5'', H-6_a), 3.93-3.83 (m, 3H, H-6'', H-8'', H-9''$_a$), 3.82-3.75 (m, 3H, H-3, H-4'', H-6_b), 3.73-3.66 (m, 2H, H-7'', H-9''$_b$), 3.53 (dd, 1H, $J_{2,1}$= 2.7 Hz, $J_{2,3}$= 11.2 Hz, H-2), 3.14 (dd, 1H, $J_{4'a,4'b}$= 17.1 Hz, $J_{4'a,5'}$= 7.5 Hz, H-4'$_a$), 2.84 (dd, 1H, $J_{4'b,4'a}$= 17.1 Hz, $J_{4'b,5'}$= 5.7 Hz, H-4'$_b$), 2.77 (dd, 1H, $J_{3''eq,3''ax}$= 13.1 Hz, $J_{3''eq,4''}$= 4.7 Hz, H-3''$_{eq}$), 2.08 (s, 3H, CH_3), 1.90-1.82 (m, 1H, H-3''$_{ax}$), **^{13}C NMR** (125 MHz, D_2O) δ: 175.7 (C_q), 175.1 (C_q), 171.7 (C_q), 167.4 (C_q), 157.6 (C_q), 99.5 (C_q), 96.3 (C-1), 94.2 (C_q), 73.5 (C-6''), 71.9 (C-5), 71.0 (C-8''), 68.4 (C-4), 67.7 (C-7''), 67.1 (C-4''), 65.2 (C-3), 63.2 (C-6), 63.0 (C-9''), 51.7 (C-5''), 51.2 (C-5'), 38.3 (C-3''), 37.9 (C-2), 30.4 (C-4'), 22.1 (CH_3)

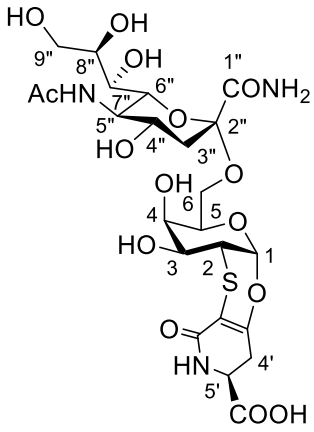
Synthesis of compound 57



Compound **55** (100 mg, 0.11 mmol) was solubilized in THF/H_2O 10:0.1 (2 mL), $Pd(OH)_2$ (20 wt.%, 10 mg) was added and the reaction was stirred at room temperature, under H_2 atmosphere. After complete conversion, the suspension was filtered through a pad of Celite$^{(R)}$, the solvents were removed and the crude was purified by flash chromatography (CH_2Cl_2/CH_3OH 8:2) to give **57** (77 mg, 86%).

ESI-MS *m/z* (%): 805.67 (100) $[M]^-$, **^{1}H NMR** (500 MHz, CD_3OD) δ: 5.69-5.62 (m, 1H, H-1), 5.45-5.37 (m, 1H, H-8''), 5.37-5.36 (m, 1H, H-7''), 4.84-4.79 (m, 1H, H-4''), 4.38- 4.31 (m, 1H, H-9''$_a$), 4.19-4.06 (m, 4H, H-5, H-5', H-6'', H-9''$_b$), 4.02-3.90 (m, 3H, H-4, H-5", H-6$_a$), 3.86 (s, 1H, OCH_3), 3.78-3.70 (m, 1H, H-3), 3.69-3.60 (m, 1H, H-6$_b$), 3.50-3.43 (m, 1H, H-2), 2.91-2.81 (m, 1H, H-4"$_a$), 2.79-2.70 (m, 1H, H-4'$_b$), 2.70-2.62 (m, 1H, H-3"$_{eq}$), 2.14 (s, 3H, CH_3), 2.12 (s, 3H, CH_3), 2.02(s, 3H, CH_3), 2.00 (s, 3H, CH_3), 1.90-1.80 (m, 4H, H-3"$_{ax}$, CH_3)

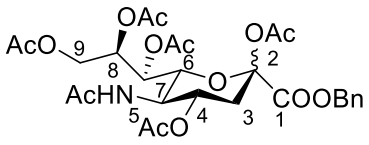
Synthesis of compound 58



Compound **57** (31 mg, 38.42 µmol) was solubilized in NH_3 in CH_3OH 4M (1 mL) and the reaction was stirred at room temperature. After 4 days, the solvent was removed and the product was purified by several washing with Et_2O to give pure **58** (16 mg, 68%).

ESI-MS *m/z* (%): 622.33 (100) $[M]^-$, **$[\alpha]^{25}_D$ =** +109.1 (c 0.001, H_2O), **^{1}H NMR** (500 MHz, D_2O) δ: 5.81 (d, 1H, $J_{1,2}$= 2.7 Hz, H-1), 4.34-4.29 (m, 1H, H-5), 4.17-4.12 (m, 1H, H-5'), 4.08-4.05 (m, 1H, H-4), 4.01-3.73 (m, 8H, H-3, H-4'', H-5'', H-6_a, H-6_b, H-6'', H-8'', H-$9''_a$), 3.72-3.62 (m, 2H, H-7'', H-$9''_b$), 3.51 (dd, 1H, $J_{2,1}$= 2.7 Hz, $J_{2,3}$= 11.2 Hz, H-2), 2.96 (dd, 1H, $J_{4'a,4'b}$= 16.9 Hz, $J_{4'a,5'}$= 6.8 Hz, H-$4'_a$), 2.81-2.71 (m, 2H, H-$4'_b$,H-$3''_{eq}$), 2.05 (s, 3H, CH_3), 1.88-1.80 (m, 1H, H-$3''_{ax}$), **^{13}C NMR** (125 MHz, D_2O) δ: 177.3 (C_q), 175.0 (C_q), 171.7 (C_q), 167.5 (C_q), 158.2 (C_q), 99.5 (C_q), 96.1 (C-1), 93.9 (C_q), 73.4 (C-6''), 71.8 (C-5), 71.0 (C-8''), 68.4 (C-4), 67.7 (C-7''), 67.1 (C-4''), 65.1 (C-3), 63.2 (C-6), 63.0 (C-9''), 52.9 (C-5''), 51.6 (C-5'), 38.2 (C-3''), 38.1 (C-2), 31.2 (C-4'), 22.0 (CH_3)

Synthesis of compound 59

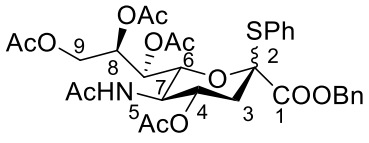


To a suspension of N-acetylneuraminic acid (5.00 g, 16.16 mmol) in DMF (25 mL) cooled to - 0 °C, DBU (3.91 g, 25.72 mmol) and BnBr (4.40 g, 25.72 mmol) were added and the mixture was stirred at room temperature overnight. After complete conversion, the solvent was removed *in vacuo.* The crude was dissolved in pyridine/Ac_2O 2:1 (100 mL)

and a catalytical amount of DMAP was added. After overnight, the reaction was diluted with CH_2Cl_2 and washed with HCl 1M (x5). The organic layer was dried over Na_2SO_4 and concentrated under vacuum. The crude was purified by flash chromatography (CH_2Cl_2/CH_3OH 9:1) to give **59** (7.39 g, 75%, β anomer > 90%).

^{1}H NMR (β anomer) (500 MHz, $CDCl_3$) δ: 7.41-7.33 (m, 5H, Bn), 5.42-5.33 (m, 2H, NH, H-7), 5.31-5.16 (m, 3H, H-4, CH_2Bn), 5.13-5.08 (m, 1H, H-8), 4.46 (dd, 1H, $J_{9a,9b}$= 12.6 Hz, $J_{9a,8}$= 2.6 Hz, H-9_a), 4.21-4.07 (m, 3H, H-5, H-6, H-9_b), 2.57 (dd, 1H, $J_{3eq,3ax}$= 13.6 Hz, $J_{3eq,4}$= 5.2 Hz, H-3_{eq}), 2.14 (s, 3H, CH_3), 2.13-2.07 (m, 4H, H-3_{ax}, CH_3), 2.05-2.03 (s, 9H, CH_3, CH_3, CH_3), 1.91 (s, 3H, CH_3), **^{13}C NMR (β anomer)** (125 MHz, $CDCl_3$) δ: 171.0 (C_q), 170.5 (C_q), 170.3 (C_q), 170.2 (2C_q), 168.3 (C_q), 165.5 (C_q), 134.9 (C_q), 128.6 (CH-Bn), 128.5 (CH-Bn), 128.3 (CH-Bn) 97.7 (C-2), 72.9 (C-6), 71.2 (C-8), 68.3 (C-4), 68.0 (CH_2-Bn), 67.8 (C-7), 62.0 (C-9), 49.4 (C-5), 35.8 (C-3), 23.2 (CH_3), 20.9 (CH_3), 20.8 (3CH_3), 20.7 (CH_3).

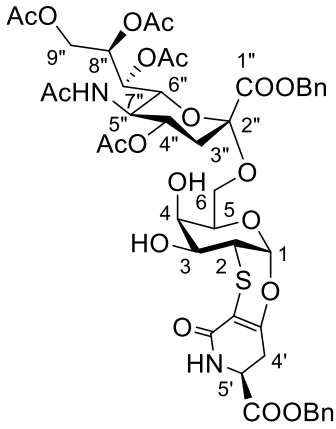
Synthesis of compound 60



To a solution of **59** (2.20 g, 3.61 mmol) in CH_2Cl_2 dry (25 mL), cooled to 0 °C, PhSH (432 mg, 3.92 mmol) and $BF_3{\cdot}Et_2O$ (580 mg – 4.09 mmol) were added and the solution was stirred at room temperature, overnight. After complete conversion, the mixture was diluted with CH_2Cl_2 and washed with $NaHCO_3$ s.s. (x5). The organic layer was dried over Na_2SO_4 and concentrated under vacuum. The crude was purified by flash chromatography (EtOAc/Hexane 8:2) to give **60** as white solid (1.97 g, 83%, β anomer > 90%).

^{1}H NMR (β anomer) (500 MHz, $CDCl_3$) δ: 7.38-7.16 (m, 10H, SPh, Bn), 6.10 (d, 1H, $J_{NH,5}$= 10.3 Hz NH), 5.47-5.46 (m, 1H, H-7), 5.44-5.38 (m, 1H, H-4), 5.08-5.05 (m, 2H, CH_2Bn), 4.90-4.87 (m, 1H, H-8), 4.62 (dd, 1H, $J_{6,5}$= 10.5 Hz, $J_{6,7}$= 2.7 Hz, H-6), 4.38 (dd, 1H, $J_{9a,9b}$= 12.3 Hz, $J_{9a,8}$= 2.4 Hz, H-9a), 4.19-4.09 (m, 2H, H-5, H-9b), 2.65 (dd, 1H, $J_{3eq,3ax}$= 13.8 Hz, $J_{3eq,4}$= 4.8 Hz, H-3_{eq}), 2.17-2.06 (m, 4H, H-3_{ax}, CH_3), 2.03 (s, 3H, CH_3), 2.03 (s, 3H, CH_3), 1.94 (s, 3H, CH_3), 1.88 (s, 3H, CH_3), **^{13}C NMR (β anomer)** (125 MHz, $CDCl_3$) δ: 171.2, 171.0, 170.4, 170.2, 170.2, 167.4, 135.8, 134.9, 129.6, 129.1, 128.7, 128.7, 128.5, 128.5, 88.5 (C-2), 73.1 (C-8, C-6), 69.2 (C-4), 68.8 (C-7), 67.6 (CH_2Bn), 62.2 (C-9), 49.1 (C-5), 37.5 (C-3), 23.1 (CH_3), 21.0 (CH_3), 20.9 (CH_3), 20.8 (CH_3), 20.7 (CH_3).

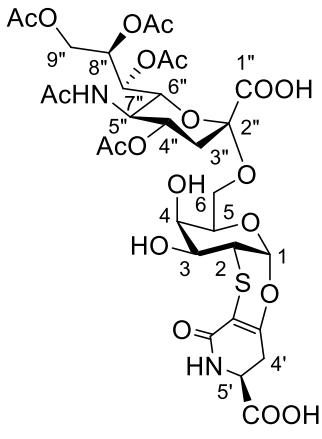
Synthesis of compound 61



To a suspension of **50** (134 mg, 0.29 mmol) and **60** (460 mg, 0.69 mmol) in CH_3CN/CH_2Cl_2 dry 10:1 (3 mL) cooled to - 40 °C, NIS (260 mg, 0.29 mmol) and TfOH (86 mg, 0.57 mmol) were added and the solution was stirred at - 40 °C under an N_2 atmosphere. After 2 h, the reaction was quenched with Et_3N (pH 8), diluted with CH_2Cl_2 and washed with $Na_2S_2O_3$ 1 M (x3). The organic layer was dried over Na_2SO_4 and concentrated under vacuum. The crude was solubilized in AcOH 80% in H_2O (10 mL) and the reaction was stirred overnight at 40 °C. The solvent was removed *in vacuo* and the product was purified by flash chromatography (EtOAc/CH_3OH 98:2) to give pure **61** (183 mg, 65% over 2 steps)

$[\alpha]^{24}_D$ = +112.3 (c 0.001, $CHCl_3$), **^{1}H NMR (α anomer)** (500 MHz, $CDCl_3$) δ: 7.42-7.34 (m, 5H, Bn), 6.21 (bs, 1H, NH), 5.61 (d, 1H, $J_{NH,5''}$= 9.8 Hz, NH), 5.61 (d, 1H, $J_{1,2}$= 2.7 Hz, H-1), 5.40-5.28 (m, 3H, NH, H-7'', H-8''), 5.27-5.19 (m, 4H, CH_2Bn, CH_2Bn), 4.93-4.84 (m, 1H, H-4''), 4.40 (dd, 1H, $J_{9''a,9''b}$= 12.6 Hz, $J_{9a,8}$= 2.6 Hz, H-9''$_a$), 4.36-4.30 (m, 1H, H-5'), 4.18-3.93 (m, 6H, H-4, H-5, H-5'', H-6$_a$, H-6'', H-9''$_b$), 3.73-3.63 (m, 2H, H-3, H-6$_b$), 3.53 (dd, 1H, 1H, $J_{2,1}$= 2.7 Hz, $J_{2,3}$= 10.7 Hz, H-2), 3.44 (bs, 1H, OH), 3.15 (bs, 1H, OH), 2.95-2.86 (m, 1H, H-4"$_a$), 2.85-2.75 (m, 1H, H-4'$_b$), 2.66 (dd, 1H, $J_{3''eq,3''ax}$= 12.8 Hz, $J_{3''eq,4''}$= 4.6 Hz, H-3''$_{eq}$), 2.15 (s, 3H, CH_3), 2.14 (s, 3H, CH_3), 2.05-2.03 (m, 7H, H-3"$_{ax}$,2CH_3), 1.89 (s, 3H, CH_3), **^{13}C NMR (α anomer)** (125 MHz, $CDCl_3$) δ:171.1, 170.8, 170.4, 170.3, 170.2, 169.5, 167.5, 165.1, 155.5, 134.8, 134.7, 128.8-128.4 (5C), 98.8, 96.6 (C-1), 96.2, 72.9, 71.3, 69.1 (C-8''), 68.9 (C-4''), 68.0, 67.9 (2C), 67.5 (C-7''), 65.7 (C-3), 63.5 (C-6), 62.7 (C-9''), 51.4 (C-5'), 49.2, 39.2 (C-2), 37.3 (C-3''), 30.7 (C-4'), 23.2 (CH_3), 21.1 (CH_3), 20.8-20.7 (3CH_3)

Synthesis of compound 62



Compound **61** (90 mg, 92.50 µmol) was solubilized in EtOAc/CH_3OH 1:1 (1 mL) $Pd(OH)_2$ (20 wt.%, 9 mg) was added and the reaction was stirred at room temperature, under H_2 atmosphere. After complete conversion, the suspension was filtered through a pad of Celite $^{(R)}$, the solvents were removed to give **62** (60 mg, 86%).

^{1}H NMR (500 MHz, CD_3OD) δ: 5.67 (d, 1H, $J_{1,2}$= 2.7 Hz, H-1), 5.40-5.37 (m, 1H, H-8''), 5.37-5.33 (m, 1H, H-7''), 5.04-4.96 (m, 1H, H-4''), 4.59-4.53 (m, 1H, H-6''), 4.40 (dd, 1H, $J_{9''a,9''b}$= 12.4 Hz, $J_{9a,8}$= 2.7 Hz, H-9''$_a$), 4.23-4.14 (m, 3H, H-5, H-5', H-9''$_b$), 4.06-4.03 (m, 1H, H-4), 4.00-3.89 (m, 2H, H-5'', H-6$_a$), 3.76-3.66 (m, 2H, H-3, H-6$_b$), 3.48 (dd, 1H, 1H, $J_{2,1}$= 2.9 Hz, $J_{2,3}$= 11.1 Hz, H-2), 3.00-2.92 (m, 1H, H-4''$_a$), 2.81-2.73 (m, 1H, H-4'$_b$), 2.64 (dd, 1H, $J_{3''eq,3''ax}$= 12.4 Hz, $J_{3''eq,4''}$= 4.6 Hz, H-3''$_{eq}$), 2.12 (s, 3H, CH_3), 2.11 (s, 3H, CH_3), 2.03(s, 3H, CH_3), 1.98 (s, 3H, CH_3), 1.85 (s, 3H, CH_3), 1.79-1.69 (m, 1H, H-3''$_{ax}$)

10.4.3 H1N1 Neuraminidase recognition

H1N1 inhibition assay

Enzymatic incubations were composed of substrate 2'-(4-Methylumbelliferyl)-α-D-*N*-acetylneuraminic acid (Sigma) (20 – 500 µM), compound **49** (0-1000 µM) in Tris-HCl 50mM pH 7.5, $CaCl_2$ 5 mM, NaCl 200 mM (total volume of 180µl). Reaction was started with addition of 20 µl of 0.22 mg/mL recombinant *Influenza* A virus H1N1 (R&D Systems, Bio Techne, France), and was allowed to incubate at 37°C for 10 min. Fluorescence was continuously monitored (λ_{ex} 365 nm/λ_{em} 445 nm) using a FLUOstar Omega (BMG Labtech, France). Initial rates were determined using 4-methylumbelliferone (Sigma) as standard, and kinetic constants were calculated using Prism 4.0 (GraphPad).

H1N1/ STnThr mimetic complex modeling

H1N1 neuraminidase structure bound to zanamivir (PDB 3B7E)[200] was used as a starting template for complex generation. Topology and

parameters files for compound **49** were obtained by Antechamber program using AM1-BCC charges.[144] The complex model was then immersed in a periodic water box (TIP3) and neutralized by adding Na^+ ions. This model was initially equilibrated by several cycles of minimizations (10000 steps, steepest descent) and Molecular Dynamics (50 ps, 200K). Then the protein backbone was constrained and the ligand was docked into H1N1 active site by Soft-Restrained Molecular Dynamics, using Zanamivir coordinates in 3B7E structure as template.[201] Finally, complex was equilibrated and energy minimized using minimization/dynamics cycles without restrained. All computational studies were performed using Amber ff14SB forcefield[142] with NAMD software.[202]

Part V

Conclusions

In the last few years, the paramount importance of specific antigens as target for immunotherapy has clearly been assessed. In the context of cancer vaccines design, structurally simple and broadly expressed, tumoral-associated- MUC1 antigens (Tn, TF and STn antigens) are TACAs of great interest and the development of mimetics of these widespread tumor markers has been recognized a powerful tool to overcome the intrinsic limitations of the natural counterparts.

Keeping this in mind and taking advantage from the "locked'' mimetic of the TnThr antigen (**13 - Figure 5.1.1**) recently developed, in this PhD thesis the synthesis of two new structurally constrained mimetics of MUC1 tumor-associated antigens and relevant applications of the three mimetics have been studied.

In detail:

A) A significant immunomodulatory activity has been highlighted for a breast cancer vaccine candidate obtained by conjugation of four residues of the TnThr mimetic **13** to the immunogenic protein CRM_{197} (**Figure 7.2.2.2**). The increased activation of human DCs and a reduction of tumor progression and metastasis in a triple negative breast cancer animal model, clearly showed a therapeutic efficacy of the glycoconjugate.

B) Silica nanoparticles (SiNPs) glycosylated with the TnThr mimetic **13** (GSiNPs **- Scheme 8.2.1.1**) were prepared and screened by $\{^{1}H\}$-^{13}C cross polarization (CP) under magic angle spinning (MAS) spectra. The results provided new perspectives for the use of SSNMR in the characterization at solid state of conjugated nano-vaccines and biomolecules-coated nanoparticles.

C) The synthesis of the TFThr mimetic **43 (Scheme 9.2.1.1**) was optimized in a gram-scale and relying on a combined multidisciplinary approach (NMR, X-ray, ITC), the molecular structural features that govern the crucial Gal-3/TFThr mimetic interaction were unrevealed. As expected, the chemical modification at the reducing unit of **43** did not affect the binding affinity. Indeed, the mimetic retained the binding epitope and conformation at the bound state of the natural antigen.

D) The STnThr mimetic **49** (**Scheme 10.2.1.1**) was efficiently synthesized by chemical sialylation of **13**. *In vitro* binding assays proved that **49** is recognized by N1 neuraminidase, an enzyme specific for sialyl-containing substrates, with an inhibition constant (*Ki)* of 40 ± 10 μM.

In conclusion, the results obtained with the structurally constrained mimetics herein reported, represent a significant progress in the glycomimetic field, and open the way to new future advances. Indeed, in the light of the synthesis of the new TFThr and STnThr mimetics a logical and attracting perspective is the design of novel multiantigen cancer vaccines with enhanced, polyvalent immunogenicity.

Abbreviations

(TA)MUC1 tumor-associated MUC1

[$(tBu)_2Si(OTf)_2$] di-tert-butylsilyl ditriflate

Aa amino acid

Abs antibodies

Ac_2O acetic anhydride

APCs antigen-presenting cells

BC breast cancer

BCR B cell receptors

BLI bioluminescence Imaging

Bn benzyl ester

BnBr benzyl bromide

Boc tert-Butyloxycarbonyl group

Boc_2O di-tert-butyl dicarbonate

BSA bovine serum albumin

Counts for minute cpm

CP cross polarization

CRD carbohydrate-recognition domain

CRM_{197} Cross Reactive Material 197

CTLs cytotoxic T cells

DAP 2,3-diaminopropionic acid

DBU 1,5-diazabiciclo(5.4.0)undec-7-ene

DCC N,N'-dicyclohexylcarbodiimide

DIPEA N,N-diisopropylethylamine

DMAP 4-(dimethylamino)pyridine

DMF N,N-dimethylformamide

DMSO dimethyl sulfoxide

DT Diphtheria Toxin

DXT-SCPNs dextran-based single-chain polymer nanoparticles

ECL *Erythrina cristagalli*

EDC 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide

EDTA ethylenediaminetetraacetic acid

ELISA enzyme-linked immunosorbent assay

EP petroleum ether

ESI electrospray ionization

Et_2O diethyl ether

Et_3N triethylamine

EtOAc ethyl acetate

EWG electron withdrawing group

Gal 3 Galectin-3

Gal galactose

GalNAc N-acetylgalactosamine

Glc D-Glucose

Gly glycine

GMNPs glycosyl superparamagnetic iron oxide nanoparticles

GSiNPs glycosylated silica nanoparticles

HBTU O-(benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate

HPA *Helix pomatia* agglutinin

IF immunofluorescence

IFN-γ interferon gamma

IPCF isopropyl chloroformate

isoSer isoserine

K_D dissociation constant

K_{Dapp} apparent dissociation constant

Ki inhibition constant

K_M Michaelis–Menten constants

MACS magnetic cell separation

MALDI Matrix-Assisted Laser Desorption

MD molecular dynamics

MFH magnetic fluid hyperthermia

MFI mean fluorescence intensity

MGL macrophage C-type lectin

MHC major histocompatibility complex

MI mitogenic index

MIP-3α macrophage inflammatory protein 3 alpha

MM molecular modelling

MNPs superparamagnetic iron oxide nano- particles

MAS magic angle spinning

MUC1 mucin 1

MU-NA 2'-(4-methylumbelliferyl)-alpha-D-N-acetylneuraminic acid

Neu5Ac N-acetylneuraminic acid

NHS N-hydroxysuccinimide

NIS N-iodosuccinimide

NKT natural killer T cells

NMM N-methylmorpholine

NMR nuclear magnetic resonance

NulOs non-2-ulosonic acids

OVA ovalbumin

Pam3CysSer tripalmitoyl-5-glycerylcysteinylserine

PAMPs pathogen-associated molecular patterns

PBMC peripheral blood adherent mononuclear cells

PBMC peripheral blood monocytes

PFA paraformaldehyde

PhtNSCl phthalimidesulfenyl chloride

Pro proline

PRRs pattern recognition receptors

***p*TsOH** *p*-toluenesulfonic acid

Py pyridine

RAFT regioselectively addressable functionalized template

RMSD root-mean-square deviation

ROI region of interest

SASA solvent accessible surface area

SAXS small-angle X-ray scattering

SD standard deviation

Ser serine

SINPs silica nanoparticles

SM3 anti-MUC1 antibody

SO_2Cl_2 sulfuryl chloride

SPR surface plasmon resonance

SSNMR solid-state NMR

STD saturation transfer difference

TAAs tumor-associated antigens

TACAs tumor-associated carbohydrate antigens

TAMs M2-polarized tumor associated macrophages

TAs tumor antigens

TBAHF tetrabutylammonium hydrogen fluoride

TBTU 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate

Tcc T cell clones

TCR T cell receptors

TD thymus-dependent

TFA trifluoroacetic acid

TfOH trifluoromethanesulfonic acid

TGFβ1 transforming growth factor beta 1

T_H helper T cells

THF tetrahydrofuran

Thr threonine

TI thymus-independent

TILs tumor-infiltrating lymphocytes

TLR Toll-like receptors

T_M memory T cells

TMB 3,3',5,5'-tetramethylbenzidine

Tn218 anti-Tn antibody

TNBC triple negative BC

TNBS 2,4,6-Trinitrobenzene sulfonic acid

TNF-α tumor necrosis factor alpha

T_{REG} regulatory T cells

TSAs tumor-specific antigens

VAA *viscumin album* agglutinin

VNTR variable number tandem repeat

WHO World Health Organization

Bibliography

1. Hollingsworth, M. A. & Swanson, B. J. Mucins in cancer: protection and control of the cell surface. *Nat. Rev. Cancer* **4**, 45–60 (2004).
2. Springer, G. T and Tn, general carcinoma autoantigens. *Science (80-.).* **224**, 1198–1206 (1984).
3. Chentoufi, A. A., Nesburn, A. B. & Benmohamed, L. Recent advances in multivalent self adjuvanting glycolipopeptide vaccine strategies against breast cancer. *Arch. Immunol. Ther. Exp. (Warsz).* **57**, 409–423 (2009).
4. Bermejo, I. A. *et al.* Water Sculpts the Distinctive Shapes and Dynamics of the Tumor-Associated Carbohydrate Tn Antigens: Implications for Their Molecular Recognition. *J. Am. Chem. Soc.* **140**, 9952–9960 (2018).
5. Nativi, C., Papi, F. & Roelens, S. Tn antigen analogues: the synthetic way to "upgrade" an attracting tumour associated carbohydrate antigen (TACA). *Chem. Commun.* **55**, 7729–7736 (2019).
6. Abbas, K. A., Lichtman, H. A. & Pillai, S. *Cellular and Molecular Immunology*. (Elsevier, 2017).
7. Wild, D., John, R., Sheehan, C., Binder, S. & He, J. *The Immunoassay Handbook Theory and applications of ligand binding, ELISA and related techniques*. (Elsevier, 2013).
8. Chaplin, D. D. Overview of the immune response. *J. Allergy Clin. Immunol.* **125**, S3–S23 (2010).
9. Nicholson, L. B. The immune system. *Essays Biochem.* **60**, 275–301 (2016).
10. Xu, M. M., Pu, Y., Zhang, Y. & Fu, Y. X. The Role of Adaptive Immunity in the Efficacy of Targeted Cancer Therapies. *Trends Immunol.* **37**, 141–153 (2016).
11. Alberts, D. & Hess, M. L. *Fundamentals of Cancer Prevention*. (Springer, 2013).
12. Spurrell, E. L. & Lockley, M. Adaptive immunity in cancer immunology and therapeutics. *Ecancermedicalscience* **8**, 1–10 (2014).
13. Zitvogel, L., Tesniere, A. & Kroemer, G. Cancer despite immunosurveillance: immunoselection and immunosubversion. *Nat. Rev. Immunol.* **6**, 715–727 (2006).
14. Polonsky, M., Chain, B. & Friedman, N. Clonal expansion under the microscope: studying lymphocyte activation and differentiation using live-cell imaging. *Immunol. Cell Biol.* **94**, 242–249 (2016).

15. Roitt, I. & Delves, P. *Encyclopedia of Immunology*. (Academic Press, 1998).

16. Parker, D. C. T Cell-Dependent B Cell Activation. *Annu. Rev. Immunol.* **11**, 331–360 (1993).

17. Mond, J. J., Lees, A. & Snapper, C. M. T Cell-Independent Antigens Type 2. *Annu. Rev. Immunol.* **13**, 655–692 (1995).

18. Obukhanych, T. V. & Nussenzweig, M. C. T-independent type II immune responses generate memory B cells. *J. Exp. Med.* **203**, 305–310 (2006).

19. Yuseff, M. I., Pierobon, P., Reversat, A. & Lennon-Duménil, A. M. How B cells capture, process and present antigens: A crucial role for cell polarity. *Nat. Rev. Immunol.* **13**, 475–486 (2013).

20. Murphy, K. & Weaver, C. *Janeway's Immunobiology*. (Garland Science, 2017).

21. Groscurth, P. & Filgueira, L. Killing Mechanisms of Cytotoxic T Lymphocytes. *Physiology* **13**, 17–21 (1998).

22. Mosmann, T. R., Li, L. & Sad, S. Functions of CD8 T-cell subsets secreting different cytokine patterns. *Semin. Immunol.* **9**, 87–92 (1997).

23. Kondělková, K. *et al.* Regulatory T cells (Treg) and Their Roles in Immune System with Respect to Immunopathological Disorders. *Acta Medica (Hradec Kral. Czech Republic)* **53**, 73–77 (2010).

24. Ariotti, S., Haanen, J. B. & Schumacher, T. N. Behavior and Function of Tissue-Resident Memory T cells. in *Advances in Immunology* vol. 114 203–216 (Elsevier Inc., 2012).

25. Jameson, S. C., Lee, Y. J. & Hogquist, K. A. Innate Memory T cells. in *Advances in Immunology* vol. 126 173–213 (Elsevier Inc., 2015).

26. Watson, D. L. Antigen Presentation. in *Immunology Guidebook* vol. 35 267–276 (Elsevier, 2004).

27. Neefjes, J., Jongsma, M. L. M., Paul, P. & Bakke, O. Towards a systems understanding of MHC class i and MHC class II antigen presentation. *Nat. Rev. Immunol.* **11**, 823–836 (2011).

28. Avery, O. T. & Goebel, W. F. CHEMO-IMMUNOLOGICAL STUDIES ON CONJUGATED CARBOHYDRATE-PROTEINS. *J. Exp. Med.* **50**, 533–550 (1929).

29. Anderson, P. W. *et al.* Effect of oligosaccharide chain length, exposed terminal group, and hapten loading on the antibody response of human adults and infants to vaccines consisting of Haemophilus influenzae type b capsular antigen unterminally coupled to the diphtheria protein. *J. Immunol.* **142**, 2464–2468 (1989).

30. Micoli, F. *et al.* Glycoconjugate vaccines: current approaches towards faster vaccine design. *Expert Rev. Vaccines* **18**, 881–895 (2019).

31. McNutt, M. Cancer Immunotherapy. *Science (80-.).* **342**, 1417–1417 (2013).

32. Maciejko, L., Smalley, M. & Goldman, A. Cancer Immunotherapy and Personalized Medicine: Emerging Technologies and Biomarker Based Approaches. *J. Mol. Biomark. Diagn.* **08**, 139–148 (2017).

33. Galluzzi, L. *et al.* Classification of current anticancer immunotherapies. *Oncotarget* **5**, 66 (2014).

34. Fan, Y. *et al.* Progress of immune checkpoint therapy in the clinic (Review). *Oncol. Rep.* **41**, 3–14 (2019).

35. Dine, J., Gordon, R., Shames, Y., Kasler, M. & Barton-Burke, M. Immune checkpoint inhibitors: An innovation in immunotherapy for the treatment and management of patients with cancer. *Asia-Pacific J. Oncol. Nurs.* **4**, 127–135 (2017).

36. Darvin, P., Toor, S. M., Sasidharan Nair, V. & Elkord, E. Immune checkpoint inhibitors: recent progress and potential biomarkers. *Exp. Mol. Med.* **50**, 1–11 (2018).

37. Alatrash, G., Crain, A. K. & Molldrem, J. J. Tumor-Associated Antigens. in *Immune Biology of Allogeneic Hematopoietic Stem Cell Transplantation* 107–125 (Elsevier, 2019). doi:10.1016/B978-0-12-812630-1.00007-4.

38. Valilou, S. F. & Rezaei, N. Tumor Antigens. in *Vaccines for Cancer Immunotherapy* 61–74 (Elsevier, 2019). doi:10.1016/B978-0-12-814039-0.00004-7.

39. Springer, G. F. Immunoreactive T and Tn epitopes in cancer diagnosis, prognosis, and immunotherapy. *J. Mol. Med.* **75**, 594–602 (1997).

40. Meeusen, E., Lim, E. & Mathivanan, S. Secreted Tumor Antigens – Immune Biomarkers for Diagnosis and Therapy. *Proteomics* **17**, 1–7 (2017).

41. Parmiani, G., De Filippo, A., Novellino, L. & Castelli, C. Unique Human Tumor Antigens: Immunobiology and Use in Clinical Trials. *J. Immunol.* **178**, 1975–1979 (2007).

42. Parmiani, G. *et al.* Universal and Stemness-Related Tumor Antigens: Potential Use in Cancer Immunotherapy. *Clin. Cancer Res.* **13**, 5675–5679 (2007).

43. Curigliano, G., Locatelli, M., Fumagalli, L. & Goldhirsch, A. Immunizing against breast cancer: A new swing for an old sword. *The Breast* **18**, S51–

S54 (2009).

44. DeMaria, P. J. & Bilusic, M. Cancer Vaccines. *Hematol. Oncol. Clin. North Am.* **33**, 199–214 (2019).

45. Buonaguro, L., Petrizzo, A., Tornesello, M. L. & Buonaguro, F. M. Translating tumor antigens into cancer vaccines. *Clin. Vaccine Immunol.* **18**, 23–34 (2011).

46. Lollini, P.-L., Cavallo, F., Nanni, P. & Forni, G. Vaccines for tumour prevention. *Nat. Rev. Cancer* **6**, 204–216 (2006).

47. Scanlon, C. S. & D'Silva, N. J. Personalized medicine for cancer therapy. *Oncoimmunology* **2**, e23433 (2013).

48. Ghazarian, H., Idoni, B. & Oppenheimer, S. B. A glycobiology review: Carbohydrates, lectins and implications in cancer therapeutics. *Acta Histochem.* **113**, 236–247 (2011).

49. Beatty, P. L., Narayanan, S., Gariepy, J., Ranganathan, S. & Finn, O. J. Vaccine against MUC1 Antigen Expressed in Inflammatory Bowel Disease and Cancer Lessens Colonic Inflammation and Prevents Progression to Colitis-Associated Colon Cancer. *Cancer Prev. Res.* **3**, 438–446 (2010).

50. Swallow, D. M. *et al.* The human tumour-associated epithelial mucins are coded by an expressed hypervariable gene locus PUM. *Nature* **328**, 82–84 (1987).

51. Singh, R. & Bandyopadhyay, D. MUC1: A target molecule for cancer therapy. *Cancer Biol. Ther.* **6**, 481–486 (2007).

52. Wei, M. M., Wang, Y. S. & Ye, X. S. Carbohydrate-based vaccines for oncotherapy. *Med. Res. Rev.* **38**, 1003–1026 (2018).

53. Stroopinsky, D., Kufe, D. & Avigan, D. MUC1 in hematological malignancies. *Leuk. Lymphoma* **57**, 2489–2498 (2016).

54. Brossart, P. *et al.* The epithelial tumor antigen MUC1 is expressed in hematological malignancies and is recognized by MUC1-specific cytotoxic T-lymphocytes. *Cancer Res.* **61**, 6846–6850 (2001).

55. Martínez-Sáez, N., Peregrina, J. M. & Corzana, F. Principles of mucin structure: Implications for the rational design of cancer vaccines derived from MUC1-glycopeptides. *Chem. Soc. Rev.* **46**, 7154–7175 (2017).

56. Gaidzik, N., Westerlind, U. & Kunz, H. The development of synthetic antitumour vaccines from mucin glycopeptide antigens. *Chem. Soc. Rev.* **42**, 4421–4442 (2013).

57. Dausset, J., Moullec, J. & Bernard J. Acquired hemolytic anemia with polyagglutinability of red blood cells due to a new factor present in

normal human serum (Anti-Tn). *Blood* **14**, 1079–93 (1959).

58. Howard, D. R. & Taylor, C. R. An Antitumor Antibody in Normal Human Serum. *Oncology* **37**, 142–148 (1980).

59. Anglin, J. H., Lerner, M. P. & Nordquist, R. E. Blood group-like activity released by human mammary carcinoma cells in culture. *Nature* **269**, 254–255 (1977).

60. Laurent, J. C., Noël, P. & Faucon, M. Expression of a cryptic cell surface antigen in primary cell cultures from human breast cancer. *Biomedicine* **29**, 260–1 (1978).

61. Jeschke, U. *et al.* Binding of galectin-1 (gal-1) to the Thomsen–Friedenreich (TF) antigen on trophoblast cells and inhibition of proliferation of trophoblast tumor cells in vitro by gal-1 or an anti-TF antibody. *Histochem. Cell Biol.* **126**, 437–444 (2006).

62. Fu, C. *et al.* Tumor-associated antigens: Tn antigen, sTn antigen, and T antigen. *HLA* **88**, 275–286 (2016).

63. Munkley, J. The role of sialyl-Tn in cancer. *Int. J. Mol. Sci.* **17**, (2016).

64. Fernández-Tejada, A., Cañada, F. J. & Jiménez-Barbero, J. Recent Developments in Synthetic Carbohydrate-Based Diagnostics, Vaccines, and Therapeutics. *Chem. - A Eur. J.* **21**, 10616–10628 (2015).

65. Videira, P. A. *et al.* Vaccine and Antibody Therapy Against Thomsen-Friedenreich Tumor-Associated Carbohydrate Antigens. *Recent Adv. Biotechnol.* **3**, 3–38 (2016).

66. Wilson, R. M. & Danishefsky, S. J. A vision for vaccines built from fully synthetic tumor-associated antigens: From the laboratory to the clinic. *J. Am. Chem. Soc.* **135**, 14462–14472 (2013).

67. Urban, D., Skrydstrup, T., Beau, J. & Skrydstrup, T. First synthesis of a C-glycoside anologue of a tumor-associated carbohydrate antigen employing samarium diiodide promoted C-glycosylation. *Chem. Commun.* **0**, 955–956 (1998).

68. Röhrig, C. H., Takhi, M. & Schmidt, R. R. Synthesis of a C-Glycoside Analog of the Tumor-Associated Tn Antigen. *Synlett* **2001**, 1170–1172 (2001).

69. Cipolla, L., La Ferla, B., Lay, L., Peri, F. & Nicotra, F. Stereoselective synthesis of α-C-glycosides of N-acetylgalactosamine. *Tetrahedron Asymmetry* **11**, 295–303 (2000).

70. Peri, F., Cipolla, L., Rescigno, M., La Ferla, B. & Nicotra, F. Synthesis and biological evaluation of an anticancer vaccine containing the C-glycoside analogue of the Tn epitope. *Bioconjug. Chem.* **12**, 325–328 (2001).

71. Bousquet, E. *et al.* Synthesis and immunostimulating activity of a thioglycolipopeptide glycomimetic as a potential anticancer vaccine derived from Tn antigen1. *J. Carbohydr. Chem.* **19**, 527–541 (2000).

72. Geraci, C. *et al.* Calix[4]arene decorated with four Tn antigen glycomimetic units and P 3CS immunoadjuvant: Synthesis, characterization, and anticancer immunological evaluation. *Bioconjug. Chem.* **19**, 751–758 (2008).

73. Companon, I. *et al.* Structure-Based Design of Potent Tumor-Associated Antigens: Modulation of Peptide Presentation by Single-Atom O/S or O/Se Substitutions at the Glycosidic Linkage. *J. Am. Chem. Soc.* **141**, 4063–4072 (2019).

74. Miller, N., Williams, G. M. & Brimble, M. A. Synthesis of fish antifreeze neoglycopeptides using microwave-assisted 'click chemistry'. *Org. Lett.* **11**, 2409–2412 (2009).

75. Grigalevicius, S. *et al.* Chemoselective Assembly and Immunological Evaluation of Multiepitopic Glycoconjugates Bearing Clustered Tn Antigen as Synthetic Anticancer Vaccines. *Bioconjug. Chem.* **16**, 1149–1159 (2005).

76. Galan, M. C., Dumy, P. & Renaudet, O. Multivalent glyco(cyclo)peptides. *Chem. Soc. Rev.* **42**, 4599–4612 (2013).

77. Pifferi, C., Berthet, N. & Renaudet, O. Cyclopeptide scaffolds in carbohydrate-based synthetic vaccines. *Biomater. Sci.* **5**, 953–965 (2017).

78. Corzana, F. *et al.* Rational design of a Tn antigen mimic. *Chem. Commun.* **47**, 5319–5321 (2011).

79. Martínez-Sáez, N. *et al.* Mucin architecture behind the immune response: design, evaluation and conformational analysis of an antitumor vaccine derived from an unnatural MUC1 fragment. *Chem. Sci.* **7**, 2294–2301 (2016).

80. Tovillas, P. *et al.* Tn Antigen Mimics by Ring-Opening of Chiral Cyclic Sulfamidates with Carbohydrate C1- S - and C1- O -Nucleophiles. *J. Org. Chem.* **83**, 4973–4980 (2018).

81. Venturi, F. *et al.* A New Scaffold for the Stereoselective Synthesis of α- O -Linked Glycopeptide Mimetics. *J. Org. Chem.* **69**, 6153–6155 (2004).

82. Jiménez-Barbero, J. *et al.* α- O -Linked Glycopeptide Mimetics: Synthesis, Conformation Analysis, and Interactions with Viscumin, a Galactoside-Binding Model Lectin. *Chem. - A Eur. J.* **15**, 10423–10431 (2009).

83. Fernández, E. M. S. *et al.* Tn Antigen Mimics Based on sp2-Iminosugars

with Affinity for an anti-MUC1 Antibody. *Org. Lett.* **18**, 3890–3893 (2016).

84. Grinstead, J. S., Koganty, R. R., Krantz, M. J., Longenecker, B. M. & Campbell, A. P. Effect of Glycosylation on MUC1 Humoral Immune Recognition: NMR Studies of MUC1 Glycopeptide−Antibody Interactions. *Biochemistry* **41**, 9946–9961 (2002).

85. Kinarsky, L. *et al.* Conformational studies on the MUC1 tandem repeat glycopeptides: Implication for the enzymatic O-glycosylation of the mucin protein core. *Glycobiology* **13**, 929–939 (2003).

86. Corzana, F. *et al.* Serine versus Threonine Glycosylation: The Methyl Group Causes a Drastic Alteration on the Carbohydrate Orientation and on the Surrounding Water Shell. *J. Am. Chem. Soc.* **129**, 9458–9467 (2007).

87. Madariaga, D. *et al.* Serine versus threonine glycosylation with α-o-GalNac: Unexpected selectivity in their molecular recognition with lectins. *Chem. - A Eur. J.* **20**, 12616–12627 (2014).

88. Martínez-Sáez, N. *et al.* Deciphering the Non-Equivalence of Serine and Threonine O-Glycosylation Points: Implications for Molecular Recognition of the Tn Antigen by an anti-MUC1 Antibody. *Angew. Chemie - Int. Ed.* **54**, 9830–9834 (2015).

89. Coltart, D. M. *et al.* Principles of mucin architecture: Structural studies on synthetic glycopeptides bearing clustered mono-, di-, tri-, and hexasaccharide glycodomains. *J. Am. Chem. Soc.* **124**, 9833–9844 (2002).

90. Madariaga, D. *et al.* Detection of tumor-associated glycopeptides by Lectins: The peptide context modulates carbohydrate recognition. *ACS Chem. Biol.* **10**, 747–756 (2015).

91. Movahedin, M. *et al.* Glycosylation of MUC1 influences the binding of a therapeutic antibody by altering the conformational equilibrium of the antigen. *Glycobiology* **27**, 677–87 (2017).

92. Capozzi, G., Menichetti, S., Nativi, C., Rosi, A. & Valle, G. Phthalimidosulfenyl chloride. Part 5. Reaction with enolizable carbonyl compounds and synthesis of functionalized thiones. *Tetrahedron* **48**, 9023–9032 (1992).

93. Capozzi, G., Fragai, M., Menichetti, S. & Nativi, C. Regio- and Stereoselective Ene and Tandem “Ene-Cycloaddition” Reactions of 2,4-Dioxopentane-3-thione. *European J. Org. Chem.* **1999**, 3375–3379 (1999).

94. Ardá, A. *et al.* Structural Insights into the Binding of Sugar Receptors (Lectins) to a Synthetic Tricyclic Tn Mimetic and Its Glycopeptide Version. *European J. Org. Chem.* **2015**, 6823–6831 (2015).

95. Lee, Y. C. & Lee, R. T. Carbohydrate-Protein Interactions: Basis of Glycobiology. *Acc. Chem. Res.* **28**, 321–327 (1995).

96. Lundquist, J. J. & Toone, E. J. The Cluster Glycoside Effect. *Chem. Rev.* **102**, 555–578 (2002).

97. Cecioni, S., Imberty, A. & Vidal, S. Glycomimetics versus Multivalent Glycoconjugates for the Design of High Affinity Lectin Ligands. *Chem. Rev.* **115**, 525–561 (2015).

98. Manuelli, M. *et al.* Iron oxide superparamagnetic nanoparticles conjugated with a conformationally blocked alfa-Tn antigen mimetic for macrophage activation. *Nanoscale* **6**, 7643–7655 (2014).

99. Richichi, B. *et al.* A Cancer Therapeutic Vaccine based on Clustered Tn-Antigen Mimetics Induces Strong Antibody-Mediated Protective Immunity. *Angew. Chemie Int. Ed.* **53**, 11917–11920 (2014).

100. Gracia, R. *et al.* Biocompatible single-chain polymer nanoparticles loaded with an antigen mimetic as potential anticancer vaccine. *ACS Macro Lett.* **7**, 196–200 (2018).

101. Jordan, A., Scholz, R., Wust, P., Fähling, H. & Roland Felix. Magnetic fluid hyperthermia (MFH): Cancer treatment with AC magnetic field induced excitation of biocompatible superparamagnetic nanoparticles. *J. Magn. Magn. Mater.* **201**, 413–419 (1999).

102. Schrand, A. M. *et al.* Metal-based nanoparticles and their toxicity assessment. *Wiley Interdiscip. Rev. Nanomedicine Nanobiotechnology* **2**, 544–568 (2010).

103. Yokoigawa, N. *et al.* Enhanced production of interleukin 6 in peripheral blood monocytes stimulated with mucins secreted into the bloodstream. *Clin. Cancer Res.* **11**, 6127–6132 (2005).

104. Bray, F. & Soerjomataram, I. *The Changing Global Burden of Cancer: Transitions in Human Development and Implications for Cancer Prevention and Control*. *Cancer: Disease Control Priorities, Third Edition (Volume 3)* vol. 49 (2015).

105. Floudas, C. S., Brar, G. & Greten, T. F. Immunotherapy: Current Status and Future Perspectives. *Dig. Dis. Sci.* **64**, 1030–1040 (2019).

106. Hollingsworth, R. E. & Jansen, K. Turning the corner on therapeutic cancer vaccines. *npj Vaccines* **4**, 7 (2019).

107. Danishefsky, S. J., Shue, Y.-K., Chang, M. N. & Wong, C.-H. Development of Globo-H cancer vaccine. *Acc. Chem. Res.* **48**, 643–52 (2015).

108. Buskas, T., Li, Y. & Boons, G.-J. The Immunogenicity of the Tumor-

Associated Antigen Lewisy May Be Suppressed by a Bifunctional Cross-Linker Required for Coupling to a Carrier Protein. *Chem. - A Eur. J.* **10**, 3517–3524 (2004).

109. Gilewski, T. *et al.* Immunization of metastatic breast cancer patients with a fully synthetic globo H conjugate: A phase I trial. *Proc. Natl. Acad. Sci.* **98**, 3270–3275 (2001).

110. Helling, F. *et al.* GM2-KLH Conjugate Vaccine: Increased Immunogenicity in Melanoma Patients after Administration with Immunological Adjuvant QS-21. *Cancer Res.* **55**, 2783–2788 (1995).

111. Livingston, P. O. *et al.* Improved survival in stage III melanoma patients with GM2 antibodies: a randomized trial of adjuvant vaccination with GM2 ganglioside. *J. Clin. Oncol.* **12**, 1036–1044 (1994).

112. Osinaga, E. *et al.* Development of an immuno-lectin-enzymatic assay for the detection of serum cancer-associated glycoproteins bearing Tn determinant. *Int. J. Oncol.* **8**, 401–406 (1996).

113. Cai, H. *et al.* Variation of the Glycosylation Pattern in MUC1 Glycopeptide BSA Vaccines and Its Influence on the Immune Response. *Angew. Chemie Int. Ed.* **51**, 1719–1723 (2012).

114. Renaudet, O., BenMohamed, L., Dasgupta, G., Bettahi, I. & Dumy, P. Towards a Self-Adjuvanting Multivalent B and T cell Epitope Containing Synthetic Glycolipopeptide Cancer Vaccine. *ChemMedChem* **3**, 737–741 (2008).

115. Lakshminarayanan, V. *et al.* Immune recognition of tumor-associated mucin MUC1 is achieved by a fully synthetic aberrantly glycosylated MUC1 tripartite vaccine. *Proc. Natl. Acad. Sci.* **109**, 261–266 (2012).

116. Yin, X.-G. *et al.* IgG Antibody Response Elicited by a Fully Synthetic Two-Component Carbohydrate-Based Cancer Vaccine Candidate with α-Galactosylceramide as Built-in Adjuvant. *Org. Lett.* **19**, 456–459 (2017).

117. Toyokuni, T. *et al.* Synthetic vaccines: synthesis of a dimeric Tn antigen-lipopeptide conjugate that elicits immune responses against Tn-expressing glycoproteins. *J. Am. Chem. Soc.* **116**, 395–396 (1994).

118. Li, Q. & Guo, Z. Recent Advances in Toll Like Receptor-Targeting Glycoconjugate Vaccines. *Molecules* **23**, 1583 (2018).

119. van Duin, D., Medzhitov, R. & Shaw, A. C. Triggering TLR signaling in vaccination. *Trends Immunol.* **27**, 49–55 (2006).

120. Malito, E. *et al.* Structural basis for lack of toxicity of the diphtheria toxin mutant CRM197. *Proc. Natl. Acad. Sci.* **109**, 5229–5234 (2012).

121. Shinefield, H. R. Overview of the development and current use of CRM197 conjugate vaccines for pediatric use. *Vaccine* **28**, 4335–4339 (2010).

122. Pecetta, S. *et al.* Carrier priming effect of CRM 197 is related to an enhanced B and T cell activation in meningococcal serogroup A conjugate vaccination. Immunological comparison between CRM 197 and diphtheria toxoid. *Vaccine* **34**, 2334–2341 (2016).

123. Kamboj, K. K., King, C. L., Greenspan, N. S., Kirchner, H. L. & Schreiber, J. R. Immunization with Haemophilus influenzae Type b–CRM 197 Conjugate Vaccine Elicits a Mixed Th1 and Th2 CD4 + T Cell Cytokine Response That Correlates with the Isotype of Antipolysaccharide Antibody. *J. Infect. Dis.* **184**, 931–935 (2001).

124. Stein, K. E. Thymus-Independent and Thymus-Dependent Responses to Polysaccharide Antigens. *J. Infect. Dis.* **165**, S49–S52 (1992).

125. Tsuchiya, A. *et al.* Prognostic relevance of tn expression in Breast Cancer. *Breast Cancer* **6**, 175–180 (1999).

126. Koboldt, D. C. *et al.* Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).

127. Richichi, B. *et al.* GM-3 Lactone Mimetic Interacts with CD4 and HIV-1 Env Proteins, Hampering HIV-1 Infection without Inducing a Histopathological Alteration. *ACS Infect. Dis.* **2**, 564–571 (2016).

128. Wua, A. M., Wub, J. H., Kuoa, H.-W. & Herpa, A. Further characterization of the binding properties of two monoclonal antibodies recognizing human Tn red blood cells. *J. Biomed. Sci.* **12**, 153–166 (2005).

129. Gibadullin, R., Farnsworth, D. W., Barchi, J. J. & Gildersleeve, J. C. GalNAc-Tyrosine Is a Ligand of Plant Lectins, Antibodies, and Human and Murine Macrophage Galactose-Type Lectins. *ACS Chem. Biol.* **12**, 2172–2182 (2017).

130. Tao, K., Fang, M., Alroy, J. & Sahagian, G. G. Imagable 4T1 model for the study of late stage breast cancer. *BMC Cancer* **8**, 228 (2008).

131. Solatycka, A. *et al.* MUC1 in human and murine mammary carcinoma cells decreases the expression of core 2 1,6-N-acetylglucosaminyltransferase and -galactoside 2,3-sialyltransferase. *Glycobiology* **22**, 1042–1054 (2012).

132. Stergiou, N., Glaffig, M., Jonuleit, H., Schmitt, E. & Kunz, H. Immunization with a Synthetic Human MUC1 Glycopeptide Vaccine against Tumor-Associated MUC1 Breaks Tolerance in Human MUC1 Transgenic Mice. *ChemMedChem* **12**, 1424–1428 (2017).

133. Virgilio, A. *et al.* Novel pyrimidopyrimidine derivatives for inhibition of cellular proliferation and motility induced by h-prune in breast cancer. *Eur. J. Med. Chem.* **57**, 41–50 (2012).

134. Zollo, M. *et al.* Overexpression of h-prune in breast cancer is correlated with advanced disease status. *Clin. Cancer Res.* **11**, 199–205 (2005).

135. D'Angelo, A. *et al.* Prune cAMP phosphodiesterase binds nm23-H1 and promotes cancer metastasis. *Cancer Cell* **5**, 137–149 (2004).

136. Freire, T., Lo-Man, R., Bay, S. & Leclerc, C. Tn Glycosylation of the MUC6 Protein Modulates Its Immunogenicity and Promotes the Induction of Th17-biased T Cell Responses. *J. Biol. Chem.* **286**, 7797–7811 (2011).

137. Kryczek, I. *et al.* Phenotype, distribution, generation, and functional and clinical relevance of Th17 cells in the human tumor environments. *Blood* **114**, 1141–1149 (2009).

138. Amedei, A. *et al.* Ex vivo analysis of pancreatic cancer-infiltrating T lymphocytes reveals that ENO-specific Tregs accumulate in tumor tissue and inhibit Th1/Th17 effector cell functions. *Cancer Immunol. Immunother.* **62**, 1249–1260 (2013).

139. Pernot, P. *et al.* Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **20**, 660–664 (2013).

140. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS : a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282 (2003).

141. Guinier, A. La diffraction des rayons X aux très petits angles : application à l'étude de phénomènes ultramicroscopiques. *Ann. Phys. (Paris).* **11**, 161–237 (1939).

142. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).

143. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **25**, 1157–1174 (2004).

144. Jakalian, A., Jack, D. B. & Bayly, C. I. Fast, Efficient Generation of High-Quality atomic charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **23**, 1623–1641 (2002).

145. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N ·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

146. Spano, D. & Zollo, M. Tumor microenvironment: a main actor in the metastasis process. *Clin. Exp. Metastasis* **29**, 381–395 (2012).

147. Win, W. W. & Franck, R. W. Model Steroid Glycoside Synthesis via a Glycosyl Transfer Mediated by Heterocycloaddition. *J. Org. Chem.* **62**, 4510–4512 (1997).

148. Homma, M., Takei, Y., Murata, A., Inoue, T. & Takeoka, S. A ratiometric fluorescent molecular probe for visualization of mitochondrial temperature in living cells. *Chem. Commun.* **51**, 6194–6197 (2015).

149. Postma, T. *et al.* Dynamic Combinatorial Chemistry with Novel Dithiol Building Blocks: Towards New Structurally Diverse and Adaptive Screening Collections. *Synlett* **24**, 765–769 (2013).

150. Müller, C., Despras, G. & Lindhorst, T. K. Organizing multivalency in carbohydrate recognition. *Chem. Soc. Rev.* **45**, 3275–3302 (2016).

151. Marradi, M., Chiodo, F., García, I. & Penadés, S. Glyconanoparticles as multifunctional and multimodal carbohydrate systems. *Chem. Soc. Rev.* **42**, 4728–4745 (2013).

152. Mourdikoudis, S., Pallares, R. M. & Thanh, N. T. K. Characterization techniques for nanoparticles: comparison and complementarity upon studying nanoparticle properties. *Nanoscale* **10**, 12871–12934 (2018).

153. Fragai, M. *et al.* SSNMR of biosilica-entrapped enzymes permits an easy assessment of preservation of native conformation in atomic detail. *Chem. Commun.* **50**, 421–423 (2014).

154. Ravera, E. *et al.* Biosilica-Entrapped Enzymes Studied by Using Dynamic Nuclear-Polarization-Enhanced High-Field NMR Spectroscopy. *ChemPhysChem* **16**, 2751–2754 (2015).

155. Martelli, T. *et al.* Atomic-Level Quality Assessment of Enzymes Encapsulated in Bioinspired Silica. *Chem. - A Eur. J.* **22**, 425–432 (2016).

156. Ravera, E. *et al.* 1H-detected solid-state NMR of proteins entrapped in bioinspired silica: a new tool for biomaterials characterization. *Sci. Rep.* **6**, 27851 (2016).

157. Giuntini, S., Cerofolini, L., Ravera, E., Fragai, M. & Luchinat, C. Atomic structural details of a protein grafted onto gold nanoparticles. *Sci. Rep.* **7**, 17934 (2017).

158. Ravera, E. *et al.* Solid-State NMR of PEGylated Proteins. *Angew. Chemie Int. Ed.* **55**, 2446–2449 (2016).

159. Cerofolini, L. *et al.* Characterization of PEGylated Asparaginase: New Opportunities from NMR Analysis of Large PEGylated Therapeutics.

Chem. – A Eur. J. **25**, 1984–1991 (2019).

160. Papi, F. *et al.* Nanoparticles for the multivalent presentation of a TnThr mimetic and as tool for solid state NMR coating investigation. *Pure Appl. Chem.* **91**, 1471–1478 (2019).

161. Sapsford, K. E. *et al.* Functionalizing Nanoparticles with Biological Molecules: Developing Chemistries that Facilitate Nanotechnology. *Chem. Rev.* **113**, 1904–2074 (2013).

162. Helle, M. *et al.* Surface Chemistry Architecture of Silica Nanoparticles Determine the Efficiency of in Vivo Fluorescence Lymph Node Mapping. *ACS Nano* **7**, 8645–8657 (2013).

163. Iltis, I. *et al.* In vivo detection of the effects of preconditioning on LNCaP tumors by a TNF-α nanoparticle construct using MRI. *NMR Biomed.* **27**, 1063–1069 (2014).

164. Touisni, N. *et al.* Fluorescent Silica Nanoparticles with Multivalent Inhibitory Effects towards Carbonic Anhydrases. *Chem. - A Eur. J.* **21**, 10306–10309 (2015).

165. Hall, P. R. *et al.* Multivalent presentation of antihantavirus peptides on nanoparticles enhances infection blockade. *Antimicrob. Agents Chemother.* **52**, 2079–2088 (2008).

166. Mahalingam, V. *et al.* Directed Self-Assembly of Functionalized Silica Nanoparticles on Molecular Printboards through Multivalent Supramolecular Interactions. *Langmuir* **20**, 11756–11762 (2004).

167. Chen, L. *et al.* The toxicity of silica nanoparticles to the immune system. *Nanomedicine* **13**, 1939–1962 (2018).

168. Alessandrini *et al.* Surface modifications of silica nanoparticles are crucial for their inert versus proinflammatory and immunomodulatory properties. *Int. J. Nanomedicine* **9**, 2815 (2014).

169. Bian, C.-F., Zhang, Y., Sun, H., Li, D.-F. & Wang, D.-C. Structural Basis for Distinct Binding Properties of the Human Galectins to Thomsen-Friedenreich Antigen. *PLoS One* **6**, e25007 (2011).

170. Ahmad, N. *et al.* Thermodynamic binding studies of cell surface carbohydrate epitopes to galectins-1, -3, and -7: Evidence for differential binding specificities. *Can. J. Chem.* **80**, 1096–1104 (2002).

171. Gupta, G. S., Gupta, A. & Gupta, R. K. *Animal lectins: Form, function and clinical applications*. *Animal Lectins: Form, Function and Clinical Applications* vol. 9783709110 (2012).

172. Mori, Y. *et al.* Binding of Galectin-3, a β-Galactoside-binding Lectin, to

MUC1 Protein Enhances Phosphorylation of Extracellular Signal-regulated Kinase 1/2 (ERK1/2) and Akt, Promoting Tumor Cell Malignancy. *J. Biol. Chem.* **290**, 26125–26140 (2015).

173. Glinskii, O. V. *et al.* Inhibition of prostate cancer bone metastasis by synthetic TF antigen mimic/galectin-3 inhibitor lactulose-L-leucine. *Neoplasia* **14**, 65–73 (2012).

174. Santarsia, S. *et al.* Molecular Recognition of a Thomsen–Friedenreich Antigen Mimetic Targeting Human Galectin-3. *ChemMedChem* **13**, 2030–2036 (2018).

175. Schmidt, R. R. & Michel, J. Facile Synthesis of α- and β-O-Glycosyl Imidates; Preparation of Glycosides and Disaccharides. *Angew. Chemie Int. Ed. English* **19**, 731–732 (1980).

176. Diniz, A., Dias, J. S., Jiménez-Barbero, J., Marcelo, F. & Cabrita, E. J. Protein-Glycan Quinary Interactions in Crowding Environment Unveiled by NMR Spectroscopy. *Chem. - A Eur. J.* **23**, 13213–13220 (2017).

177. Schumann, F. H. *et al.* Combined chemical shift changes and amino acid specific chemical shift mapping of protein–protein interactions. *J. Biomol. NMR* **39**, 275–289 (2007).

178. Yongye, A. B. *et al.* Molecular Recognition of the Thomsen-Friedenreich Antigen–Threonine Conjugate by Adhesion/Growth Regulatory Galectin-3: Nuclear Magnetic Resonance Studies and Molecular Dynamics Simulations. *Biochemistry* **51**, 7278–7289 (2012).

179. Rodriguez, M. C. *et al.* Thermodynamic Switch in Binding of Adhesion/Growth Regulatory Human Galectin-3 to Tumor-Associated TF Antigen (CD176) and MUC1 Glycopeptides. *Biochemistry* **54**, 4462–4474 (2015).

180. Fielding, L. NMR methods for the determination of protein-ligand dissociation constants. *Prog. Nucl. Magn. Reson. Spectrosc.* **51**, 219–242 (2007).

181. Binder, F. P. C., Lemme, K., Preston, R. C. & Ernst, B. Sialyl-Lewisx - ein "präorganisiertes Wasseroligomer"? *Angew. Chemie* **124**, 7440–7444 (2012).

182. Knirel, Y. A., Shashkov, A. S., Tsvetkov, Y. E., Jansson, P.-E. & Zähringer, U. 5,7-DIAMINO-3,5,7,9-TETRADEOXYNON-2-ULOSONIC ACIDS IN BACTERIAL GLYCOPOLYMERS: CHEMISTRY AND BIOCHEMISTRY. in *Advances in Carbohydrate Chemistry and Biochemistry* vol. 58 371–417 (2003).

183. Vimr, E. R., Kalivoda, K. A., Deszo, E. L. & Steenbergen, S. M. Diversity of

Microbial Sialic Acid Metabolism. *Microbiol. Mol. Biol. Rev.* **68**, 132–153 (2004).

184. Schoenhofen, I. C., Vinogradov, E., Whitfield, D. M., Brisson, J.-R. & Logan, S. M. The CMP-legionaminic acid pathway in Campylobacter: Biosynthesis involving novel GDP-linked precursors. *Glycobiology* **19**, 715–725 (2009).

185. Chen, X. & Varki, A. Advances in the Biology and Chemistry of Sialic Acids. *ACS Chem. Biol.* **5**, 163–176 (2010).

186. Passaniti, A. & Hart, G. W. Cell surface sialylation and tumor metastasis. Metastatic potential of B16 melanoma variants correlates with their relative numbers of specific penultimate oligosaccharide structures. *J. Biol. Chem.* **263**, 7591–603 (1988).

187. Schauer, R. *Sialic Acids*. vol. 10 (Springer Vienna, 1982).

188. Crocker, P. R., Paulson, J. C. & Varki, A. Siglecs and their roles in the immune system. *Nat. Rev. Immunol.* **7**, 255–266 (2007).

189. De Meo, C. & Jones, B. T. *Chemical Synthesis of Glycosides of N-Acetylneuraminic Acid*. *Advances in Carbohydrate Chemistry and Biochemistry* vol. 75 (Elsevier Inc., 2018).

190. Prendergast, J. M. *et al.* Novel anti-Sialyl-Tn monoclonal antibodies and antibody-drug conjugates demonstrate tumor specificity and anti-tumor activity. *MAbs* **9**, 615–627 (2017).

191. Loureiro, L. R. *et al.* Novel monoclonal antibody L2A5 specifically targeting sialyl-Tn and short glycans terminated by alpha-2–6 sialic acids. *Sci. Rep.* **8**, 12196 (2018).

192. Tanaka-Okamoto, M. *et al.* Identification of internally sialylated carbohydrate tumor marker candidates, including Sda/CAD antigens, by focused glycomic analyses utilizing the substrate specificity of neuraminidase. *Glycobiology* **28**, 247–260 (2018).

193. Marra, A. & Sinay, P. Acetylation of N-acetylneuraminic acid and its methyl ester. *Carbohydr. Res.* **190**, 317–322 (1989).

194. Marra, A. & Sinaÿ, P. Stereoselective synthesis of 2-thioglycosides of N-acetylneuraminic acid. *Carbohydr. Res.* **187**, 35–42 (1989).

195. Ranade, S. C. & Demchenko, A. V. Mechanism of chemical glycosylation: Focus on the mode of activation and departure of anomeric leaving groups. *J. Carbohydr. Chem.* **32**, 1–43 (2013).

196. Air, G. M. Influenza neuraminidase. *Influenza Other Respi. Viruses* **6**, 245–256 (2012).

197. Krueger, A. C. *et al.* Synthesis of potent pyrrolidine influenza

neuraminidase inhibitors. *Bioorg. Med. Chem. Lett.* **18**, 1692–1695 (2008).

198. Prevato, M. *et al.* Expression and Characterization of Recombinant, Tetrameric and Enzymatically Active Influenza Neuraminidase for the Setup of an Enzyme-Linked Lectin-Based Assay. *PLoS One* **10**, e0135474 (2015).

199. Lloren, K. K. S. *et al.* In Vitro and In Vivo Characterization of Novel Neuraminidase Substitutions in Influenza A(H1N1)pdm09 Virus Identified Using Laninamivir-Mediated In Vitro Selection. *J. Virol.* **93**, e01825-18 (2019).

200. Xu, X., Zhu, X., Dwek, R. A., Stevens, J. & Wilson, I. A. Structural Characterization of the 1918 Influenza Virus H1N1 Neuraminidase. *J. Virol.* **82**, 10493–10501 (2008).

201. Lafite, P., André, F., Zeldin, D. C., Dansette, P. M. & Mansuy, D. Unusual Regioselectivity and Active Site Topology of Human Cytochrome P450 2J2 †. *Biochemistry* **46**, 10237–10247 (2007).

202. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).