

# **Coiled-Coil Peptides as Multivalent Scaffolds for Carbohydrates:**

from receptor targeting to vaccines exploiting sugar-protein interactions



## **Dissertation**

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Department of Biology, Chemistry, Pharmacy

Institute of Chemistry and Biochemistry

Freie Universität Berlin

**Elsa Zacco**

from Reggio Calabria, Italy

**July 2015**



1<sup>st</sup> Reviewer: **Prof. Dr. Beate Koksch**

2<sup>nd</sup> Reviewer: **Prof. Dr. Peter H. Seeberger**

Defense date: 25.09.2015



*Al Coraggio. E alla Donna che mi ha insegnato cos'è.*



## DECLARATION

This Ph.D. thesis was carried out from June 2011 until July 2015 in the research group of Prof. Dr. Beate Koksch at the Institute for Chemistry and Biochemistry, Faculty of Chemistry, Biology and Pharmacy of the Freie Universität Berlin.

I hereby declare that this PhD thesis was prepared autonomously. Content, quotes or images previously published by third parties are indicated by referring to the original work.

Additional contributions are reported below:

- TEM micrographs have been acquired by Dr. von Berlepsch at the BioSupraMol core facility of Freie Universität Berlin (Germany)
- The compounds 5-amino-pentanyl- $\alpha$ -D-mannopyranoside and 5-amino-pentanyl  $\beta$ -D-galactopyranosyl-(1 $\rightarrow$ 2)- $\alpha$ -D-mannopyranoside were provided by Dr. Christopher E. Martin and Prof. Peter H. Seeberger, from the Max Planck Institute of Colloids and Interfaces of Berlin (Germany)
- All biological and immunological assays were performed at the Max Planck Institute of Colloids and Interfaces of Berlin (Germany) with the help and the assistance of Dr. Chakkumkal Anish and Dr. Julia Hütter.

The work presented in this dissertation resulted in the following publications:

1. Tailored Presentation of Carbohydrates on a Coiled Coil-Based Scaffold for Asialoglycoprotein Receptor Targeting.

**Zacco E.**, Hütter J., Heier J.L., Mortier J., Seeberger P.H., Lepenies B. and Koksch B.  
*ACS Chem Biol.* **2015** Jun 30.

DOI: 10.1021/acscchembio.5b00435

2. A Self-Assembling Peptide Scaffold for the Multivalent Presentation of Antigens.

**Zacco E.**, Anish C., Martin C.E., von Berlepsch H., Brandenburg E., Seeberger P.H. and Koksch B.

*Biomacromolecules* **2015** Jun 4.

DOI: 10.1021/acs.biomac.5b00572

**Elsa Zacco**

Berlin, July 2015





## ACKNOWLEDGEMENTS

*“Strive not to be a success, but rather to be of value.”*

A.H.

I am truly glad to have had the chance of dedicating my time and my passion to this project. It has inspired me and taught me a lot, not only scientifically but also personally.

I would like to thank my supervisor Prof. Beate Koksich for the opportunity she has given me to work in this topic and for the faith she placed in my success. My thanks also go to all my colleagues, current and past, in particular to Allison, for her patience, support and availability, and to my friend Jason, without whom these four years would have just not been the same.

The success of this work is the result of strong and valid scientific collaborations with the group of Prof. Peter Seeberger from the Max Planck institute of Colloids and Interfaces, whose experience and advice I highly value. In particular, I had the pleasure to work alongside and be inspired by Dr. Bernd Lepenies, Dr. Chakkumkal Anish and Dr. Julia Hütter. I further thank my other collaboration partners Dr. Hans von Berlepsch, Dr. Jeremie Mortier and Dr. Christopher Martin.

Sincere thanks go also to my family and friends, who have believed in me and supported me through this stage of my life, and to my parents, who have taught me the importance of following my passions and never giving up.

Last but not least (definitively, most definitively not least), I am deeply grateful to my companion, friend, accomplice and ally Max, for the constant care, attention and unconditioned love that he donates me every day, and for being to me shelter and home.

Elsa Zacco



## ABSTRACT

Over the past decade, the employment of chemically-synthesized scaffolds for the delivery of biologically-relevant molecules has proven to be a highly advantageous strategy for medicinal purposes, primarily due to stabilization of the cargo and enhancement of site-targeted therapeutic effects. Conjugation to a scaffold has been shown to reduce ligand degradation, impart optimal geometric conformation and increase the likelihood of the active species reaching the target. A wide range of synthetic carriers, including dendrimers, nanoparticles, PNAs, LNAs and cell-penetrating peptides, has been explored; among these, peptides are of particular interest due to their diverse functional groups, folding properties, biocompatibility and low toxicity. In particular, the structural simplicity and regularity of the  $\alpha$ -helical coiled-coil folding motif makes it a suitable scaffold for multivalent ligand display. By changing only a few positions in a coiled-coil sequence, it is possible to influence the behaviour of the resulting helices, obtaining either short dimeric peptides or long, fiber-forming carriers.

This thesis explores the coiled-coil motif as a scaffold for the multivalent display of peptide and carbohydrate ligands. We aspired to demonstrate the versatility of the coiled-coil motif in two distinct projects. The first study relied on the dimeric form, providing structural predictability for the rational presentation of carbohydrates for lectin targeting. The second study aimed to increase the efficiency of carbohydrate-antibody recognition by displaying ligands on a self-assembling, fiber-forming peptide.

The first project is entitled “Tailored presentation of carbohydrate ligands on a coiled-coil scaffold for asialoglycoprotein receptor targeting”. In this work, we determined the binding of members of a coiled-coil glycopeptide library to hepatocytes and established the optimal distance and orientation of the galactose moieties for interaction with the asialoglycoprotein receptor using flow cytometry. We confirmed that binding occurs through receptor-mediated endocytosis via inhibition studies with cytochalasin D; moreover, fluorescence microscopy studies demonstrated the uptake of the carrier peptides into cells.

The second project is entitled “A self-assembling peptide scaffold for the multivalent presentation of antigens”. Here, a coiled coil-based sequence was used to create tunable higher-order structures on the nanometer scale, allowing for the multivalent presentation of a mannose moiety and a peptide epitope, the presence of which did not interfere with self-assembly of the nanostructure. The multivalent display of these ligands led to tighter binding by both mannose-specific lectins and appropriate antibodies. The potential of the novel self-assembling peptide to display antigens in bioanalytical assays that demand high sensitivity

was illustrated by decoration with a disaccharide glycotope from the surface of the *Leishmania* parasite. Anti-*Leishmania* antibodies present in human and canine sera bind their antigen more effectively in the case of multivalent display on the coiled-coil scaffold. In summary, the  $\alpha$ -helical coiled-coil scaffolds investigated here were shown to effectively multivalently present carbohydrate and/or peptide ligands to their specific binding partners, in the context of either a well-defined precision tool (project 1) or self-assembled nanofibers (project 2). Project 1 demonstrated that a coiled-coil carrier decorated with selected ligands may be tailored-made for applications involving other therapeutically-relevant receptors. Project 2 established that a synthetically accessible fiber-forming coiled-coil scaffold can provide the multivalent effect required to enhance binding avidity of specific antibodies and/or receptors.

## ZUSAMMENFASSUNG

In den letzten Jahren stellt die Anwendung von synthetischen Gerüsten für den Transport biologisch relevanter Moleküle eine vorteilhafte Strategie in der Medizin dar. Dabei wird nicht nur der Wirkstoff stabilisiert sondern auch noch zusätzlich die Wirkungsweise am Zielort moduliert. Es konnte gezeigt werden, dass durch die Konjugation einer biologisch aktiven Spezies an ein Gerüst deren Stabilität erhöht wird, die optimale geometrische Konformation vermittelt wird und generell eine Erhöhung der Wahrscheinlichkeit auf ein Erreichen des Wirkstoffes am Wirkungsort stattfindet. Zu diesem Zweck wurden unterschiedlichste synthetische Träger untersucht wie zum Beispiel: Dendrimere, Nanopartikel, PNAs, LNAs und zellpenetrierende Peptide. In diesem Kontext stehen Peptide im besonderen Fokus der Forschung aufgrund ihrer zahlreichen funktionellen Gruppen, ihres vielfältigen Faltungsverhalten und ihrer Biokompatibilität sowie geringen Toxizität. Ein prominenter Vertreter, das  $\alpha$ -helikale *coiled-coil* Faltungsmotif, ermöglicht durch seine strukturelle Einfachheit und Regelmäßigkeit die multivalente Präsentation von Liganden. Durch die Modifikation von wenigen Positionen in einem *coiled-coil* Motiv ist es möglich, einen Einfluss auf das Oligomerisierungsverhalten der resultierenden Helices auszuüben, wodurch entweder kurze Peptidimere oder lange fibrillenartige Gerüste erhalten werden.

Die vorliegende Arbeit untersucht die Möglichkeit das *coiled-coil* Motiv als ein Gerüst für die multivalente Präsentation von Kohlenhydraten oder Peptiden zu verwenden. Dazu ist es angedacht den Vorteil des *coiled-coil* Motifs in zwei unterschiedlichen Projekten zu demonstrieren. Im ersten Projekt wird die Dimer-Form genutzt, die eine strukturelle Vorhersagbarkeit besitzt um eine rationale Präsentation von Kohlenhydraten für eine Lektin-Interaktion zu ermöglichen. Die zweite Studie zielt auf eine Erhöhung der Effizienz einer Kohlenhydrat-Antikörper-Erkennung durch die Ligandenpräsentation auf einem selbstorganisierenden Fibrillen bildenden Peptid ab.

Das erste Projekt heißt “Tailored presentation of carbohydrate ligands on a coiled-coil scaffold for asialoglycoprotein receptor targeting”. In diesem Teil konnte die Bindung der Mitglieder einer *coiled-coil* Glykopeptid-Bibliothek zu Heptazyten bestimmt werden und der optimale Abstand sowie die Orientierung der Galactose-Liganden für die Wechselwirkung mit dem Asialoglykoproteinrezeptor (ASGPR) durch Durchflusszytometrie etabliert werden. Anhand von Inhibitionsstudien mit Cytochalasin D wurde bestätigt, dass die Bindung durch die ASGPR vermittelte Endozytose abläuft. Darüber hinaus wurde die Aufnahme der Trägerpeptide in die Zelle mithilfe von fluoreszenz-mikroskopischen Studien nachgewiesen.

Im zweiten Teil mit dem Namen “A self-assembling peptide scaffold for the multivalent presentation of antigens” wurde ein *coiled-coil*-basierende Sequenz genutzt um veränderbare höhergeordnete Strukturen im Nanometerbereich zu generieren. Diese ermöglichten die gleichzeitige Präsentation von Mannose-Liganden und einem Peptid-Epitop, wobei die Selbstorganisation der Helices durch die Modifikationen nicht beeinträchtigt wurde. Die multivalente Präsentation dieser Liganden resultierte

in einer stärkeren Bindung durch einerseits Mannose-bindende Lektine und andererseits durch geeignete Antikörper. Weiterhin wurde das Potential dieses neuartigen selbstorganisierenden Peptids untersucht, die eingebauten Antigene in einem bioanalytischen Assay zu präsentieren, welcher einer hohen Empfindlichkeit bedarf. Dazu wurde das Peptid mit einem Disaccharid-Glykotoptop vom *Leishmania* Parasit versehen und mit humanen und caninen anti-*Leishmania*-Antikörpern enthaltenden Seren inkubiert. Die Bindung des Antigens durch die Antikörper erfolgte im Falle der multivalenten Präsentation durch das *coiled-coil* Gerüst weitaus effektiver.

Zusammenfassend wiesen die in der vorliegenden Arbeit untersuchten  $\alpha$ -helikalen *coiled-coil* Gerüste eine effiziente multivalente Präsentation von Kohlenhydraten und/ oder Peptiden für die entsprechenden spezifisch bindenden Partner auf. Dies wurde entweder durch ein gut definiertes Präzisionswerkzeug (Projekt1) oder durch selbstorganisierenden Nanofasern (Projekt 2) erzielt. In Projekt 1 wurde dargelegt, dass *coiled-coil* Peptide, welche mit ausgewählten Liganden versehen sind, für die Anwendung zur Untersuchung von therapeutisch relevanten Rezeptoren maßgeschneidert synthetisiert werden können. Im Projekt 2 wurde gezeigt, dass synthetisch zugängliche Fibrillen bildende *coiled-coil* Gerüste einen multivalenten Effekt aufweisen, welcher für eine Verstärkung der Bindung eines spezifischen Antikörpers und/ oder Rezeptors erforderlich ist.

## ABBREVIATIONS

<b>Ab</b>	Antibody	<b>DMF</b>	Dimethylformamide
<b>Abz</b>	Aminobenzoic acid	<b>DNA</b>	Deoxyribonucleic acid
<b>ACN</b>	Acetonitrile	<b>DT</b>	Diphtheria toxin
<b>AcOH</b>	Acetone	<b>ELISA</b>	Enzyme-linked immunosorbent assay
<b>AOSP</b>	All-on-solid-phase	<b>ESI</b>	Electrospray ionization
<b>AP</b>	Alkaline phosphatase	<b>EtOAc</b>	Ethyl acetate
<b>APC</b>	Antigen presenting cells	<b>FCS</b>	Fetal calf serum
<b>ASGPR</b>	Asialoglycoprotein receptor	<b>FITC</b>	Fluorescein isothiocyanate
<b>Boc</b>	<i>tert</i> -Butyloxycarbonyl	<b>FL</b>	Fluorescence microscope
<b>BPBS</b>	PBS + 0.5% - 0.1% BSA	<b>Fmoc</b>	Fluorenylmethyloxycarbonyl
<b>BSA</b>	Bovine serum albumin	<b>FSC</b>	Forward-scattered light
<b>CCP</b>	Parent coiled-coil peptide	<b>Gal</b>	Galactose
<b>CD</b>	Circular dichroism	<b>GalNAc</b>	<i>N</i> -Acetylgalactosamine
<b>CLM</b>	Confocal laser microscopy	<b>Glc</b>	Glucose
<b>ConA</b>	Concanavalin A	<b>GlcNAc</b>	<i>N</i> -Acetylglucosamine
<b>CPP</b>	Cell-penetrating peptide	<b>GPI</b>	Glycosylphosphatidylinositol
<b>CRM197</b>	Cross-reactive material 197	<b>HATU</b>	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
<b>Cy5</b>	Cyanine 5	<b>HCl</b>	Hydrochloric acid
<b>DCM</b>	Dichloromethan	<b>HepG2</b>	Human hepatocellular carcinoma cells
<b>DIC</b>	Diisopropylcarbodiimide		
<b>DIPEA</b>	<i>N,N</i> -Diisopropylethylamine		
<b>DMEM</b>	Dulbecco's Modified Eagle Medium		

<b>Hex</b>	Hexane	<b>NT</b>	Nanotube
<b>HIV</b>	human immunodeficiency virus	<b>OD</b>	Optical density
<b>HOAt</b>	1-Hydroxy-7-azabenzotriazole	<b>PBS</b>	Phosphate buffered saline
<b>HOBt</b>	Hydroxybenzotriazole	<b>PBST</b>	PBS + Tween
<b>HPLC</b>	High-performance liquid chromatography	<b>PEG</b>	Polyethylene glycol
<b>HRP</b>	Horseradish peroxidase	<b>PEO</b>	Polyethylene oxide
<b>Ig</b>	Immunoglobulin	<b>PNA</b>	Peptide nucleic acid
<b>L</b>	Long (referred to “spacer”)	<b>PP</b>	Polyproline
<b>Lcp</b>	Left-handed circularly polarized light	<b>PTA</b>	Phosphotungstic acid
<b>LNA</b>	Locked nucleic acid	<b>PTMC</b>	Polytrimethylene carbonate
<b>LPG</b>	Lipophosphoglycan	<b>Rcp</b>	Right-handed circularly polarized light
<b>Man</b>	Mannose	<b>RP</b>	Reverse phase
<b>MeOH</b>	Methanol	<b>S</b>	Short (referred to “spacer”)
<b>MHC</b>	Major histocompatibility complex	<b>siRNA</b>	Short interfering RNA
<b>MPI</b>	Max Planck Institute	<b>SPPS</b>	Solid-phase peptide synthesis
<b>mRNA</b>	Messenger RNA	<b>SSC</b>	Side-scattered light
<b>MS</b>	Mass spectrometry	<b>TAT</b>	Trans-activating transcriptional activator
<b>Mtt</b>	<i>N</i> -Methyltrityl	<b>TBTU</b>	2-(1H-Benzotriazol-1-yl)- <i>N,N,N',N'</i> -tetramethylammonium tetrafluoroborate/hexafluorophosphate
<b>NHS</b>	<i>N</i> -hydroxyl succinimide	<b><i>t</i>Bu</b>	<i>tert</i> -Butyl
<b>NLS</b>	Nuclear localization signal		
<b>NMR</b>	Nuclear magnetic resonance		



<b>TEM</b>	Transmission electron microscopy	<b>TFA</b>	Trifluoroacetic acid
		<b>TLC</b>	Thin layer chromatography

Abbreviations of the 20 canonical amino acids are consistent with the biochemical nomenclature proposed by the IUPAC-IUB commission (*Eur J Biochem.* **1984**, 138, 9-37).



## TABLE OF CONTENT

<b>1. INTRODUCTION</b> .....	1
1.1 Peptides in medicinal chemistry.....	3
1.1.1 Pros and cons.....	4
1.1.2 Peptides as multivalent scaffolds.....	5
1.1.3 Self-assembling peptide scaffolds.....	7
1.1.4 Peptide interactions with cells.....	9
1.2 Coiled-coils in nature and as tools for medicine.....	12
1.2.1 The coiled-coil heptad repeat.....	14
1.2.2 Attaining multivalency with the coiled-coil motif.....	16
1.2.3 $\alpha$ -Helical-based scaffolds for biomedical applications.....	19
1.3 Carbohydrate-protein interactions in cells and medicine.....	21
1.3.1 Biosynthesis of glycoproteins.....	21
1.3.2 Carbohydrate functions in cells.....	23
1.3.3 Multivalency in carbohydrate-protein interactions.....	25
1.3.4 Use of carbohydrates in medicine.....	26
<b>2. APPLIED METHODS</b> .....	29
2.1 Peptide synthesis and purification.....	30
2.1.1 Solid-phase peptide synthesis.....	30
2.1.2 Reversed-phase high-performance liquid chromatography.....	31
2.1.3 Electrospray ionization-mass spectroscopy.....	31
2.2 Further analytical methods.....	32
2.2.1 Circular dichroism.....	32
2.2.2 Transmission electron microscopy.....	35
2.3 Biological methods and techniques.....	38
2.3.1 Flow cytometry.....	38
2.3.2 Confocal fluorescence microscopy.....	40
2.3.3 Enzyme-linked immunosorbent assay.....	41
2.3.4 Microdot array.....	43
2.3. Employed fluorophores.....	44
<b>3. MOTIVATION AND OBJECTIVES</b> .....	47
<b>4. PROJECT 1: Tailored presentation of carbohydrate ligands on a coiled-coil scaffold for Asialoglycoprotein receptor targeting</b> .....	49

4.1 An overview of ASGPR.....	50
4.2 Glycopeptide library design and nomenclature.....	52
4.3 Glycopeptide synthetic strategy.....	56
4.4 Glycopeptide library structural characterization.....	58
4.5 Analysis of hepatocyte-mediated uptake.....	60
4.6 Inhibition studies of ASGPR-mediated endocytosis.....	65
4.7 Analysis of the internalization of selected glycopeptides within HepG2 cells.....	66
4.8 Summary and outlook I.....	69
4.9 Experimental procedures I.....	71
<b>5. PROJECT 2: A self-assembling peptide scaffold for the multivalent presentation of antigens.....</b>	<b>79</b>
5.1 An overview of carbohydrate-based vaccines.....	80
5.1.1 The carbohydrate antigen.....	82
5.1.2 The T-cell epitope.....	83
5.1.3 The carrier molecule.....	84
5.1.4 The conjugation method.....	85
5.2 Multivalent scaffold design.....	85
5.3 Glycoconjugate synthetic strategy.....	87
5.4 Evaluation of peptide library structural features.....	90
5.4.1 Peptide library secondary structure.....	91
5.4.2 Peptide library supramolecular structure.....	94
5.5 Multivalent presentation of ligands to selected binding partners.....	98
5.5.1 Interaction of peptide epitope with anti-Diphtheria toxin antibodies.....	98
5.5.2 Interaction of the mannose ligand with lectins.....	100
5.5.3 Presentation of the antigenic disaccharide to specific antibodies.....	105
5.6 Summary and outlook II.....	107
5.7 Experimental procedures II.....	110
<b>6. REFERENCES.....</b>	<b>115</b>
<b>7. CURRICULUM VITAE.....</b>	<b>137</b>

# 1

**Synthetic Glycopeptides as Multivalent Scaffolds for Carbohydrates:**  
from receptor targeting to vaccines exploiting sugar-protein interactions

## INTRODUCTION

- 1.1 Peptides in medicinal chemistry
- 1.2 Coiled coils in nature and as tools for medicine
- 1.3 Carbohydrate-protein interactions in cells and medicine

Peptides and proteins play an essential, irreplaceable role in the dynamics and kinetics of cellular processes. It is the regulation of these particular biomolecules that allows the cell to maintain homeostasis.<sup>1, 2</sup> In the human body, they are found in every tissue and perform a wide range of essential functions that depend on the amino acid primary sequence and the 3D structure. Disclosing the molecular mechanism by which these functions are exploited and understanding how peptides interact with their biological targets may allow intervention in pathological conditions and can offer a new tool for exploiting other cellular processes. The popularity of peptides and peptidomimetics has constantly grown in the last few years, especially because of their synthetic accessibility and potential applications. Chemists and chemical biologists are particularly challenged by the tight protein structure-function correlation; the main aim is to recreate certain cellular interactions by offering synthetic peptide alternatives able to provide specific functional group arrangements that would translate into a defined tertiary structure and a specialized function. Synthetic peptides represent an appropriate simplification of a complex system, yet a system able to maintain folding properties and biological functionalities.

The helix is by far the most frequently occurring structural motif in proteins and it is associated with key biological functions such as molecular recognition, replication and pathogen invasion.<sup>3, 4</sup> It is remarkable how, in spite of their structural simplicity, helical peptides are involved in the most important cellular functions. For instance, it is the helical structure of collagen that provides mechanical strength and stability to tissues and bones;<sup>5</sup> the Lac repressor, the protein responsible for the regulation of galactose, interacts with its binding site on the *lac* operon mainly via one of its helices;<sup>6</sup> and even apoptosis can be mediated by the interaction of the helical peptides Bak and Bcl-xL with gramicidin A.<sup>7, 8</sup>

However, not all protein-protein interactions rely on direct recognition: some of these processes of molecular recognition must be finely regulated and require post-translational modifications. The most common post-translation modification is protein glycosylation. The addition of carbohydrates to proteins allows for cell-cell and cell-environment communication in both physiological and pathological processes. Examples are respectively offered by lectin-mediated processes and pathogen invasion. Lectins, such as SIGN-R1 and CSL, are carbohydrate-binding proteins involved in cell adhesion: R1 is involved in the transient recognition of neuronal cells; CSL is associated with adhesion between normal and transformed cells.<sup>9</sup> Regarding pathogen invasion, the entry mechanism of HIV is only one of many examples of pathogens relying on interactions between their own proteins and those of

the host cell to establish infection. The case of HIV represents an excellent example in which the specificity of the invasion process depends on the presence of the helical motif within glycoproteins.<sup>10</sup>

The work presented in this thesis focuses on the use of the coiled-coil helical motif for the presentation of carbohydrates to biological macromolecules for biomedical applications. This introduction describes the theoretical background and state of the art necessary for the understanding of the results of this study, and is divided as follows:

- Section 1.1, “Peptides in medicinal chemistry”, offers an overview of the use of peptides as scaffolds for biomedical applications. It defines the pros and cons of the use of peptides and describes how their self-assembly properties can be used for cell targeting;
- Section 1.2, “Coiled coils in nature and as tools for medicine”, discusses the structural characteristics of the coiled-coil helical motif and its use as a scaffold for the delivery of biologically-active molecules;
- Section 1.3, “Carbohydrate-protein interactions in cells and medicine”, underlines the biological role of carbohydrate-protein interactions and the use of these interactions in medicine.

## **1.1 PEPTIDES IN MEDICINAL CHEMISTRY**

Bioactive peptides trigger many biochemical cascades associated with cell-cell and cell-environment communication, cellular growth and death, the immune response and other modulatory activities in living systems. Therefore, it is not surprising that their employment in medicine and in the pharmaceutical industry is very rapidly increasing.

Newer generations of peptide drugs have been discovered by screening natural products or by rational design, including ligand-based, mechanism-based and receptor-based design. This class of protein-based drugs, referred as “biologics”, includes peptides such as insulin, growth factors and engineered antibody fragments. Biologics are often not suitable for oral delivery and are thus typically formulated for injection. Nevertheless, they are an extremely successful class of compounds, both economically and therapeutically. Among the top selling injectable peptides are the 10-amino acid immunomodulator Copaxone, for multiple sclerosis, the 9-amino acid gonadotropin receptor agonists Lupron and Zoladex, for breast and prostate

cancer, respectively, the 8-amino acid hormone secretion inhibitor Sandostatin, and the 2-amino acid proteasome inhibitor Velcade, for multiple myeloma.<sup>11</sup> According to a new report published by Transparency Market Research “Peptide Therapeutics Market”, the worldwide market for peptides employed in medicine, either as active ingredients or as drug adjuvants, is expected to grow to US\$ 27.3 billion by 2020,<sup>12</sup> and would likely increase further, since the production of peptides as drugs puts more emphasis in the future on personalized medicine.

### **1.1.1 Pros and cons**

The success of peptides in biomedical applications lays in their intrinsic chemical properties and biological activities:

1. Peptides are natural hormones, antibiotics, growth factors, chemokines, cytokines and neurotransmitters; they are therefore metabolically and allergenically tolerated, and this generally translates into low cell toxicity.<sup>13</sup>
2. The direct or indirect effect of peptides can be addressed towards a large variety of targets, such as infectious diseases, neurological diseases, metabolic disorders and cancer.<sup>14-16</sup> Due to their diversity in sequence and structure, their action is likely to be highly target-specific.
3. Peptides represent a valuable compromise between bulky proteins and small molecule drugs, in terms of both costs and activity. While their functionalities are available for binding, delivery and post-translational modifications just like proteins, they are typically more synthetically accessible.

As promising as drug candidates or adjuvants they may be, the process that each single peptide candidate has to face before being effectively adopted in medicine is not straight forward and several obstacles must be overcome before commercialization of a peptide drug:

1. Due to their physical and chemical properties, peptides are often poorly orally bioavailable, and, even when they do manage to overcome this limitation, they are frequently and quickly cleared by the liver and kidneys.
2. Once distributed in the blood stream, peptides typically exhibit rather short half-life, because, as proteins, they are subject to the action of several proteases.<sup>17</sup>
3. As a further obstacle, peptides tend to show very limited penetration of the blood-brain barrier, and this calls for the use of additional delivery systems.



A schematic view of the advantages and disadvantages of the use of peptides in medicinal chemistry is shown in Table 1.

ADVANTAGES	DISADVANTAGES
High activity	Low oral bioavailability
High specificity	Low stability
Low tissue accumulation	Limited solubility
Low toxicity	Short half-life
Chemical diversity	Difficult delivery

**Table 1.** Advantages and disadvantages of peptide therapeutics.

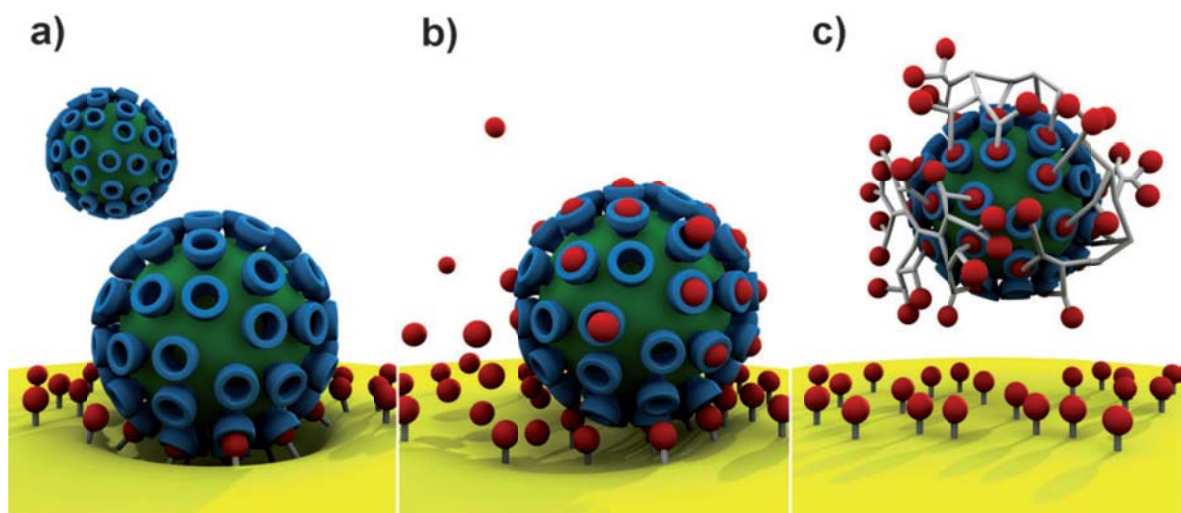
### 1.1.2 Peptides as multivalent scaffolds

To reduce the disadvantages in the use of peptides as drugs, over the past decade scaffolds have been developed to help biologically-active peptides overcome their physical and chemical limitations. These preparations are often based on nanocarrier technology, including nanoparticles, liposomes and micelles.<sup>18-20</sup> The delivery of peptides or other biologics via synthetic scaffolds offers the advantages of stabilizing the drug, improving the site-targeted therapeutic effect and enhancing efficacy and safety,<sup>21</sup> because conjugation to a stable scaffold can reduce drug degradation, offers a rigid backbone to orientate the conjugated drug in an optimal conformation and increases the chance of reaching the target with the help of specific recognition motifs.

Interestingly, peptides themselves have found wide application as biocompatible scaffolds for the delivery of drugs or other cell-interacting components (specific examples are given below). Biocompatibility of a delivery system refers to its ability to perform with an appropriate response in a specific biological environment; thus, peptides represent an ideal scaffold for targeting biological macromolecules, since the nature of their composition is perfectly compatible with the targeted system. Moreover, the selective targeting of biological macromolecules also requires a robust and structurally reliable delivery system; the advantage represented by the use of peptides in this context is that they can be custom-designed to obtain defined secondary and tertiary structures, which can be modified at selected positions with negligible structural alterations. They offer variety in terms of length, structure and functionalization. Moreover, their high surface-area-to-volume ratio and chemical diversity renders them of great interest for the delivery of biologically-relevant molecules.

Peptides represent a most versatile building block, since they can be modified and functionalized with an array of natural ligands, such as carbohydrates, nucleic acids and lipids. Their biocompatibility increases the chances of molecular recognition events with the selected target, particularly considering the specific selectivity deriving from the ability of some of them to self-assemble into ordered nanostructures.<sup>22</sup>

To rationally-design a peptide scaffold for the presentation of ligands, the most important parameter to take into account is the interaction between the displayed molecule and its specific cellular target, since the ligand-receptor interaction is characterized by high selectivity.<sup>23</sup> The basic principle of selective delivery is to increase the chance of binding and recognition events by incrementally increasing the ligand concentration at the interaction site, while lessening non-selective distribution. In the design of synthetic scaffolds, these results are generally achieved by multivalently displaying a higher number of ligands and by specifically directing the scaffold toward the target cell type/tissue. An example of multivalent ligand display is given in Figure 1.



**Figure 1.** Ligand delivery via a multivalent scaffold. **A:** A virus recognizes and binds the cell surface. **B:** The interaction virus-cell is weakly impeded by non-competitive binding with a monovalent ligand. **C:** The virus-cell interaction is efficiently shielded by multivalent display of the ligand. Adapted from Fasting *et al.* with permission of Wiley-VCH, copyright 2012.<sup>24</sup>

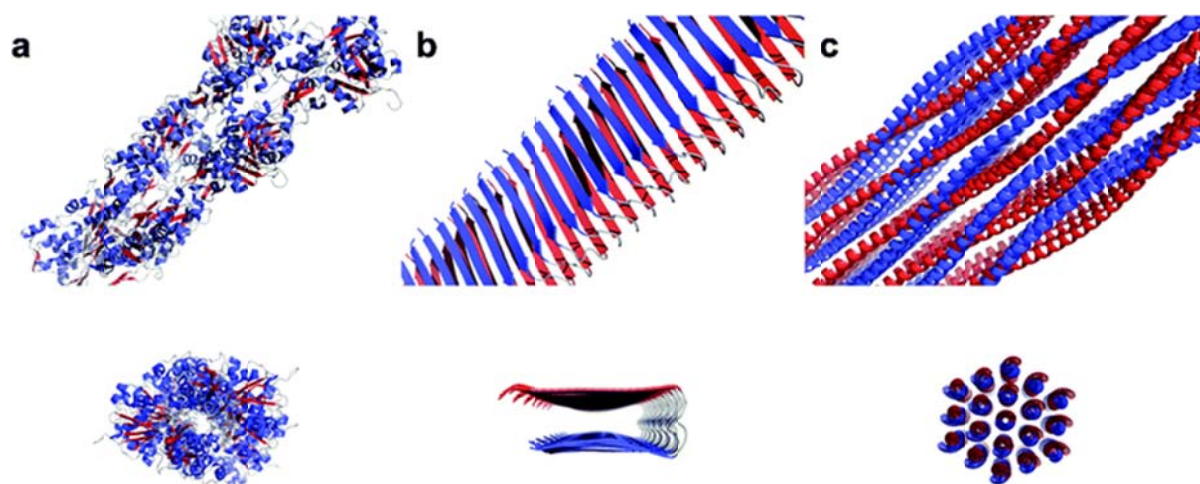
The peptide research of the past decade is full of examples of the use of peptides as biomolecular carriers. For instance, a whole cellular microenvironment has been recreated by Lutolf and Hubbell with the use a scaffold peptide reproducing the regularity and stability of the cytoskeletal network:<sup>25</sup> the nanofibrillar network was formed by the self-assembly of a

small building block presenting bioactive ligands responsive to cell-secreted signals. Other examples are offered by the presentation of neuronal and epithelial growth factors as ligands of self-assembling peptides, inducing cellular maturation and wound healing.<sup>26, 27</sup> In the first case, the Stupp group encapsulated neural progenitor cells within a three-dimensional network of peptide nanofibers presenting a neurite-promoting laminin epitope, which induces neuronal differentiation. In the second example, Schneider, Garlick and Egles created a peptide that undergoes molecular self-assembly to form a unique supramolecular structure that stably covers the surface of wounds, suggesting that this scaffold may serve as a viable wound dressing. A further use of peptides as scaffolds is represented by the cyclic peptide of Chan *et al.*,<sup>28</sup> which significantly improves the proangiogenic activity of the presented ligands at nanomolar concentrations, and is stable in human serum.

### **1.1.3 Self-assembling peptide scaffolds**

The ability of certain peptides to undergo self-assembly makes them extremely useful as tools for both medicinal chemistry and materials science. Compared to other nanoscale bio-assemblies, such as branched polymers and dendrimers, PNA, and LNA,<sup>29-31</sup> peptides offer the advantage of facile *de novo* design by fine tuning of the primary sequence, to obtain morphologies such as peptide nanotubes, amyloid fibrils and  $\alpha$ -helical nanofibers. The self-association of these types of structures is a recurring theme in both physiological and disease-associated conditions.<sup>32, 33</sup> Mimicking natural scaffolds and modifying them with the desired alterations, such as ligand conjugation, can reduce the synthetic effort and increase the chance of success, since the architectures adopted by peptide nanoassemblies are often retained in the presence of guest molecules.

During the self-assembly process, single building blocks recognize one another and associate to form two- or three-dimensional networks. Recognition is driven by a combination of noncovalent interactions, such as electrostatic and hydrophobic interactions, hydrogen bonding and aromatic stacking. Inspired by molecular recognition processes in nature, a variety of hosts of synthetic peptide nanoscaffolds have been developed. Three of the most common types of peptide structures are depicted in Figure 2 and described below.



**Figure 2.** Three examples of elongated peptide self-assemblies. **A:** Tubular assemblies deriving from a mix of secondary structures (actin). **B:** Amyloidogenic peptide A $\beta$  (1–40), solely cross- $\beta$  laminated sheets. **C:** Fibers formed by  $\alpha$ -helical coiled-coils. Adapted from Morriss and Serpell, with permission of The Royal Society of Chemistry, copyright 2010.<sup>34</sup>

- Peptide nanotubes (NTs) represent a class of protein-like self-assembly which has recently received much attention. They are defined as elongated nanoobjects with a definite inner hole and a diameter ranging from 0.7 nm to 1  $\mu$ m. NTs can be synthetically obtained from structures as simple as dipeptides<sup>35</sup> or as complex as cyclic peptides<sup>36</sup> and are often used as biomaterials. Their diversity is also reflected in the multitude of secondary structures they adopt: they can be obtained from  $\alpha$ -helical or  $\beta$ -sheet peptides, as well as peptides or proteins adopting both conformations simultaneously.
- Amyloids also derive from a spontaneous assembly of proteins or peptides into higher-ordered structures. In nature, they often associate with the so called proteinopathies, or protein misfolding diseases, such as Alzheimer's, Parkinson's and Huntington's diseases.<sup>37</sup> In a living system, they can aggregate into fibrils or ribbons and sediment in the extracellular matrix, interfering with the physical and physiological stability of cells. All amyloid fibrils share a  $\beta$ -sheet-rich secondary structure and a cross- $\beta$  sheet quaternary structure and are built up mostly by hydrophobic associations. The structural regularity and stability of amyloid peptides, together with their involvement in several human disorders, make them extremely interesting in the context of medicinal chemistry. A step in this direction is represented by the study of Kelly and coworkers on functional amyloids: the group identified a mammalian amyloidogenic protein Pmel17 which is not toxic but, on the contrary, accelerates the

polymerization of reactive small molecules into melanin, a critically important biopolymer that protects against a broad range of cytotoxic damage.<sup>38</sup> Following this trend, Maji *et al.* discovered that hormones in secretory granules of the endocrine system are stored in an amyloid-like conformation, and therefore functional amyloids can contribute to physiological cell development.

- The type of self-assembling peptide structure most-commonly found in nature is the superhelical coil, and examples of this fold are  $\alpha$ -keratin and myosin proteins.<sup>39, 40</sup> The  $\alpha$ -helical coiled-coil motif is the basic building block for the construction of fibrous filaments with lengths on the order of several hundred  $\mu\text{m}$  that bundle to form matured fibers. The chemistry of the hydrophobic and electrostatic interactions leading to the formation of the  $\alpha$ -helical fibers is well studied and understood, and the rules for the design of such supramolecular structures are known. Since only some of the amino acid residues of the coiled-coil motif are actively involved in the folding of the peptide, the remaining ones can be used for the display of ligands and functionalities. This characteristic makes the  $\alpha$ -helical fibers an optimal three-dimensional scaffold.

#### **1.1.4 Peptide interactions with cells**

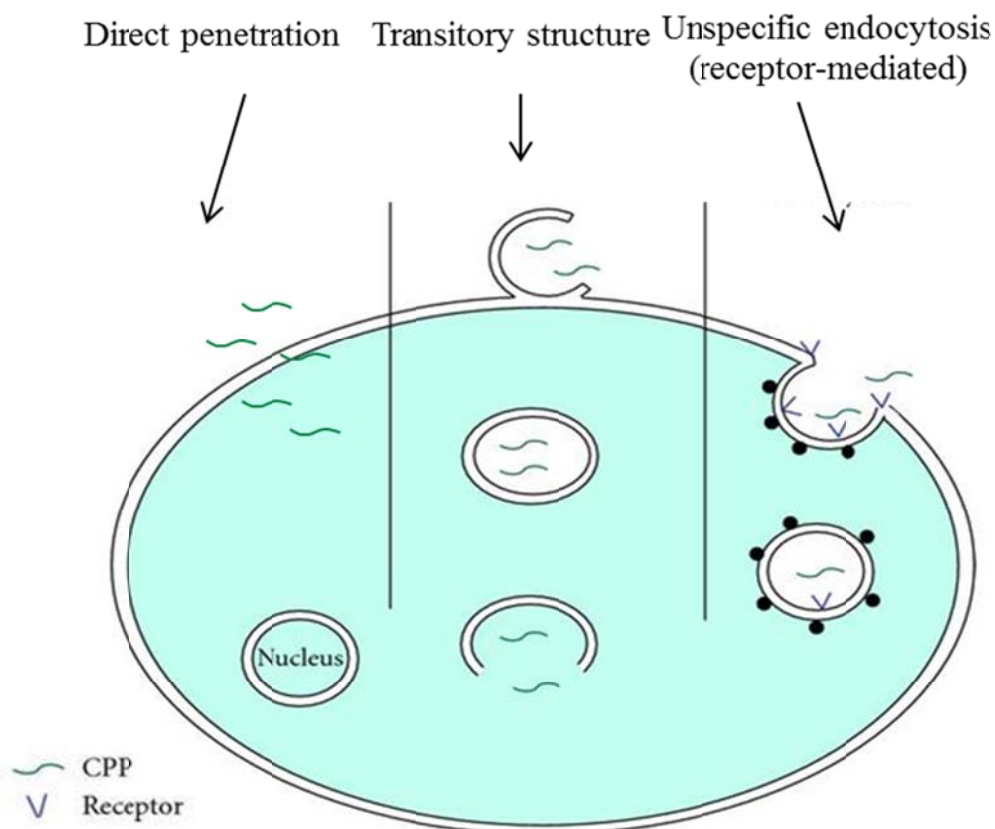
The employment of peptides as scaffolds for the delivery of drugs or biologically-relevant molecules into cells has experienced rapid growth, due to the discovery of peptides with the innate ability to cross the cell membrane, called cell-penetrating peptides (CPPs).

CPPs offer the advantage of intracellular delivery of information-rich molecules, avoiding dispersion in the blood stream and reducing side effects. On the one hand, this type of peptide shows high levels of cellular uptake and therefore has significant potential as a carrier system; on the other hand, they exhibit cellular promiscuity, to the detriment of specificity. For this reason, according to the type of molecule to be delivered and the therapeutic use of the synthetic system, peptide chemists have also turned to other peptides that do not possess cell-penetrating potential but that instead deliver their cargo inside the cell by precisely displaying a presented ligand to the cognate receptor. Thus, the CPP approach is known as passive targeting, while the ligand-receptor approach is referred to as active targeting. One current goal in this field is the optimization of peptides as delivery molecules for both selective targeting and efficient cellular internalization.

- **Passive targeting**

Passive targeting refers to the phenomenon of transport across the cell membrane without specific interactions with membrane components. In the context of peptide scaffolds, passive targeting is possible mostly in the case of CPPs. CPPs facilitate the cellular delivery of their cargo, which can be bound covalently or noncovalently. A typical CPP generally contains either a large number of surface positive charges (polycationic CPP), or alternating hydrophilic and hydrophobic residues (amphipathic CPP).

The first identified CPP was the trans-activating transcriptional activator (TAT) from HIV. Since then, a large number of CPPs have been identified and characterized.<sup>41</sup> No unique mechanism of translocation has been identified; instead, CPPs can also be classified according to their primary entrance mechanism into the cell: direct penetration, endocytosis, or the formation of transitory structures, as reported in Figure 3.



**Figure 3.** Cellular-uptake mechanism of CPPs. Representation of the three main ways of CPP internalization into cells. Adapted from Madani et al. with permission of Hindawi Publishing Corporation, copyright 2011.<sup>42</sup>

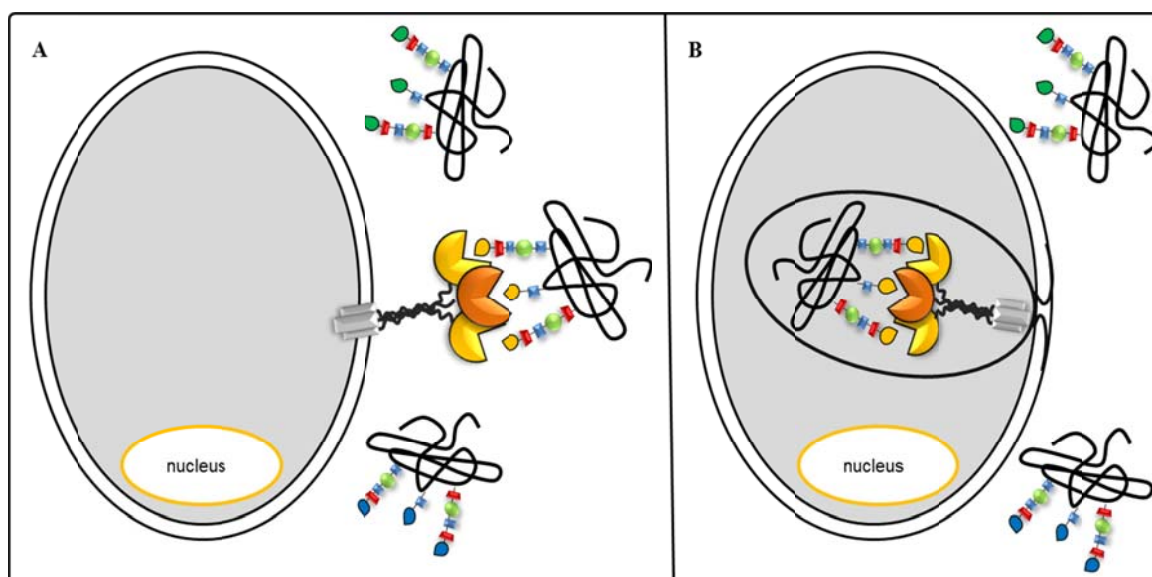
The internalization of CPPs by direct penetration is an energy-independent process and it most likely involves favorable electrostatic interactions between polycationic CPPs and negatively-charged phospholipids. In contrast, endocytosis is an energy-dependent mechanism that can occur either by nonspecific receptor mediation or by interaction between the positive charges of the CPP and surface polysaccharides of the cell (heparin sulfates). The formation of inverted micelles is the most common form of transitory structure for CPP transportation. These are formed by aggregates of colloids in combination with membrane phospholipids that allow the CPP to remain in the internal hydrophilic environment. CPPs have found numerous applications in biology and medicine as carrier systems: they have been used for the delivery of proteins, full-length antibodies, nucleic acids, small-molecule drugs and imaging agents.<sup>43-47</sup> Overall, CPPs show very high potential for further development and optimization, but their lack of specificity remains a major drawback.

- **Active targeting**

Active targeting, also called ligand-based targeting, refers to the receptor-mediated process. The main advantage of delivering cargo in this fashion is increased cell binding and internalization. Over the past decade, several peptides have been exploited for this purpose, particularly for anti-cancer therapy, anti-inflammation, diagnostics and imaging. Mai *et al.* demonstrated that a peptide selected by phage display increased the percentage of active uptake of a liposome containing an anticancer drug and the percentage of cancer cell death,<sup>48</sup> Wang and co-workers designed a peptide conjugated with PEG and PTMC as a drug carrier, which was able to be selectively target bone lesion sites and demonstrate pathology-responsive activity;<sup>49</sup> several examples exist for the use of targeting peptides carrying theranostic agents, able to simultaneously deliver a drug and a contrast agent;<sup>50</sup> Subramanian and others, for instance, developed a cationic peptide scaffold that allows its cargo to be delivered specifically into the nucleus by conjugating with a nonclassical nuclear localization signal (NLS). The system shows improved nuclear localization and DNA binding properties of the NLS.<sup>51</sup>

This active targeting mechanism is depicted in Figure 4.





**Figure 4.** Receptor-mediated active targeting. **A:** Selective binding of a carrier-ligand complex with a specific membrane receptor; only the correct binding partner can complex with its receptor. **B:** Internalization of the bound carrier-ligand complex via receptor-mediated endocytosis.

There are some limitations in the active-targeting system: the expression of the receptor must be localized to the target cells and receptor saturation frequently occurs, reducing targeting efficiency and increasing the risk of non-specific binding phenomena.

Due to their diverse functionalities and structural properties, peptides represent a scaffold for receptor-mediated targeting that can easily be fine-tuned. They can accommodate a large variety of ligands that can be placed at specific sites and precisely oriented. Moreover, by varying the length and the sequence of the peptide scaffold, these can be made to accommodate a high number of chemically distinct ligands simultaneously. These features make peptides good potential candidates for multivalent ligand presentation.

## 1.2 COILED COILS IN NATURE AND AS TOOLS FOR MEDICINE

Self-assembly plays an essential role in the production of robust and efficient nanoscaffolds. The main challenges in the synthesis of such structures are the rational design of molecules that can form defined supramolecular aggregates and the stability and toxicity of the scaffolds. Nature continuously faces these same challenges and it has, over the course of evolution, achieved absolute control over the structure and function of complex proteins. The way in which nature exerts its control is by dictating the chemical and electrical interactions



between the basic building blocks of the macromolecules, and thus by programming polypeptide chain sequences. The consequence is a tight correlation between structure and function, stability under physiological conditions and refined interactions with other macromolecules. Therefore, the rational design of peptide scaffolds must be based on natural systems to enable translating primary amino acid sequences into defined supramolecular structures. The literature offers several examples of cell-interacting peptide scaffolds. For instance, the  $\beta$ -sheet-based peptide hydrogel RAD16II, consisting of alternating hydrophilic and hydrophobic amino acids, was found to be capable of supporting the attachment and growth of a variety of mammalian cell types. The same ability to form biocompatible hydrogels is also reported for  $\beta$ -hairpin-based peptides, short aromatic peptides, hybrid amphiphilic peptides and  $\alpha$ -helical peptide hydrogels.<sup>52-55</sup> However, the state of the art regarding the incorporation of biologically-active motifs as ligands on peptide scaffolds for interaction with cells is limited to hybrid amphiphilic peptides<sup>56</sup> and  $\beta$ -sheet-based fibrils<sup>57</sup> (an overview on the subject is given by Wu, Zhang and Hauser).<sup>58</sup> To date, no coiled coil-based peptide has been reported to be employed for the delivery of biologically-relevant molecules to cells. This may be partially due to the large number of amino acids required to yield an  $\alpha$ -helical fiber, which translates into relatively high production costs at industrial levels.

Nevertheless, the  $\alpha$ -helical coiled-coil motif is a ubiquitous protein domain. Predictions based on analyses of primary sequences suggest that >5% of all proteins form coiled-coils.<sup>59</sup> It is therefore the most abundant protein fold, and it offers versatility and strength. Proteins relying on the coiled-coil motif for their interactions have a large variety of functions, their roles ranging from transcription regulation to cell architecture.<sup>60, 61</sup> Moreover, there are several proteins in which the coiled-coil motif plays a relevant pharmacological role: the SNARE proteins, involved in vesicular trafficking of neurotransmitters and viral fusion;<sup>62</sup> the GCN4 transcriptional activator protein, which binds specifically to the promoters of yeast amino acid biosynthetic genes;<sup>63</sup> the proto-oncoprotein Fos and Jun, which bind DNA via their leucine zipper motif;<sup>64</sup> and the myosin protein family, involved in the transport of specific biomolecules, vesicles and organelles in eukaryotic cells.<sup>65</sup> The crucial role played in both physiological and pathological conditions, together with the structural reliability and predictability of the coiled-coil motif, makes it a promising scaffold for biomedical and pharmaceutical applications.

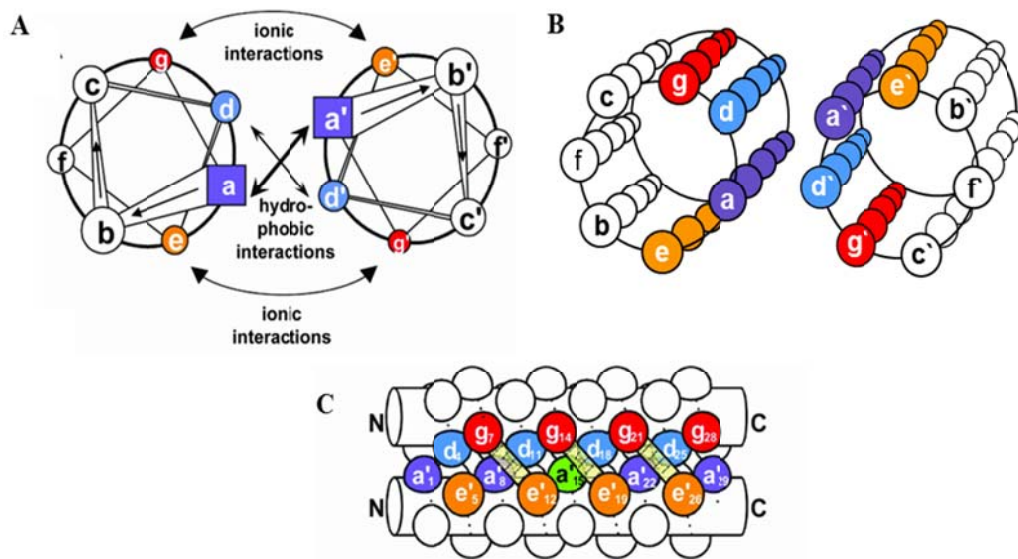
### 1.2.1 The coiled-coil heptad repeat

The  $\alpha$ -helical coiled-coil motif was identified in 1952 by Francis Crick, based on X-ray diffraction data from  $\alpha$ -keratin.<sup>66, 67</sup> According to Crick's definition, the hallmark of the coiled-coil motif is a distinctive packing of the amino acid side chains in the core of a bundle (knobs-into-holes): a residue of one helix (knob) packs into a space surrounded by the side chains of four amino acids from another helix (hole). Two or more amphipathic  $\alpha$ -helices can assemble into a supercoil (or superhelix) by wrapping themselves around each other. Most commonly, supercoils consist of two-to-five helices, running in the same (parallel) or in opposite (antiparallel) directions.<sup>68</sup> Compared to the 3.6 residues per helix turn found in 3.6<sub>13</sub> helices within globular proteins, the coiled-coil motif forces a distortion of the amino acids around each other and brings the number of residues per helix turn down to 3.5. For this reason, a complete heptad repeat is distributed every two turns of the helix.<sup>69, 70</sup>

A superhelix can be composed by either homomeric or heteromeric assemblies, according to the primary amino acid sequence, which also dictates stoichiometry, topology and assembly-disassembly equilibrium.<sup>71</sup> There are, moreover, other essential properties that influence the stability of a coiled coil, such as helical propensity of the amino acids, hydrophobicity and tightness of the core, shielding of the core from the solvent and favorable polar and ionic interactions. The contribution of each amino acid at each single location along the repeats is still under investigation, although various attempts to elucidate this process have been made.<sup>72</sup> Left-handed coiled coils are characterized by a seven-residue periodicity called heptad repeat. Right-handed coiled coils are instead defined by an eleven-residue periodicity, the undecad repeat. In the context of preparation of synthetic scaffolds, the simplicity of the heptad repeat is favored and most commonly employed.

The structural characterization of a high number of coiled coil-containing proteins by X-ray crystallography and NMR spectroscopy allows for the identification of coiled-coil basic design principles. Within the seven residues composing the coiled-coil heptad repeat, designated *a*, *b*, *c*, *d*, *e*, *f* and *g*, the main driving force for the assembly is established by placing hydrophobic residues in positions *a* and *d* to form a hydrophobic core.<sup>73</sup> Amino acids at positions *e* and *g* are often charged to stabilize the helix and give directionality.<sup>74</sup> They pack against the hydrophobic core and can participate in interhelical electrostatic interactions between residue *i* (*g* position) of one helix and residue *i*' + 5 of the other helix (*e*' position, belonging to the next heptad). Residues *b*, *c* and *f* are solvent-exposed and can therefore be used for the conjugation of ligands.<sup>75-77</sup> The positions of the heptad repeat and the

hydrophobic and electrostatic interactions of a parallel, heterodimeric coiled coil are illustrated in Figure 5.



**Figure 5.** Schematic representation of a parallel, heterodimeric coiled coil. Selected side chains are shown as balls and sticks. **A:** Helical wheel diagram indicating side chains as balls and sticks, demonstrating the hydrophobic interactions between  $a$ - $a'$  and  $d$ - $d'$  positions and the ionic interactions between  $e$ - $g'$  and  $g$ - $e'$  positions. The wheel shows a top view of the amino acid sequence. **B:** Horizontally tilted view of the helical wheel, displaying five full heptad repeats. **C:** Side view of the dimeric coiled-coil peptide, emphasizing the hydrophobic core and electrostatic interactions between  $e$ - $g'$  positions.<sup>78</sup>

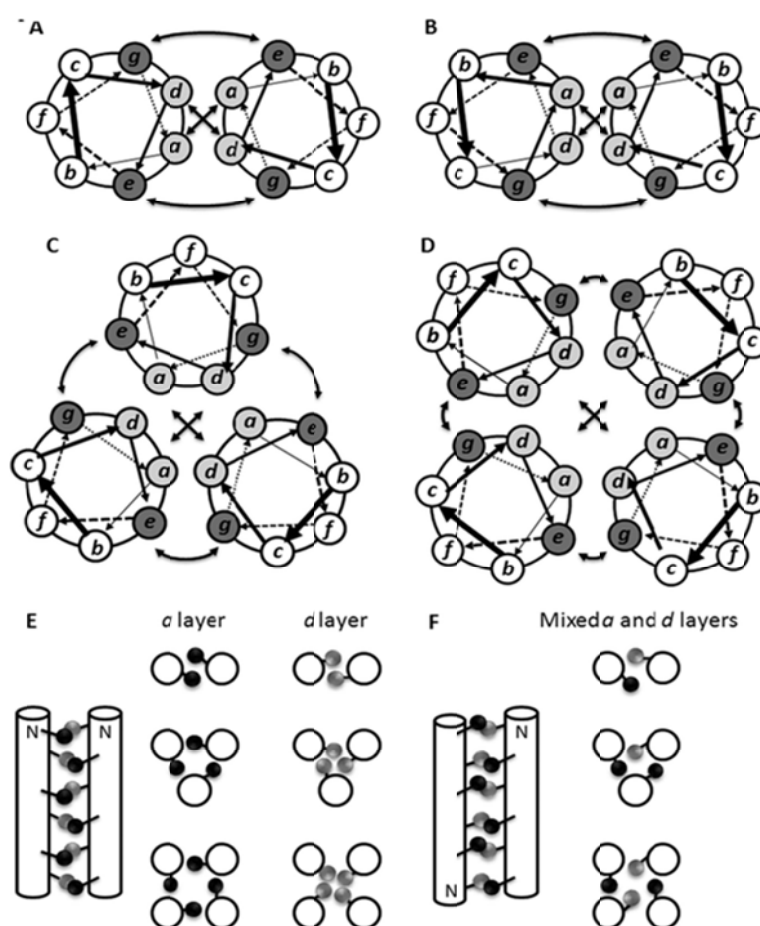
Within the hydrophobic core, the choice of amino acid to place at  $a$  and  $d$  positions have a fundamental role in stability, ellipticity and oligomerization of the coiled-coil. For example, in a parallel homodimeric coiled-coil a  $\beta$ -branched side chain such as valine packs well at the  $a$  position, while position  $d$  favors the  $\gamma$ -branching of leucine.<sup>79</sup> Placing a  $\beta$ -branched amino acid into a  $d$  position would force the parallel dimeric peptide to adopt the thermodynamically disfavored rotamer conformation and thus lead to the formation of a tetramer.<sup>80</sup> The hydrophobicity of the amino acids at the core is also the most stabilizing energetic contribution; for example, leucine at a  $d$  position contributes about  $38.5 \text{ KJ mol}^{-1}$  per coiled-coil dimer, relative to alanine.<sup>81</sup>

The main variable influencing pairing and orientation of the helices is the electrostatic interaction between the amino acid residues at  $e$  positions on one helix and at  $g'$  positions on the other helix ( $i \rightarrow i'+5$ ). These positions are commonly occupied by glutamic acid and

lysine, respectively, but in a heterologomeric system it is possible to switch the orientation of one of the helices by placing glutamic acid and lysine in *g* and *e* positions, respectively. Replacing attractive charges with repulsive charges at these positions destabilizes pairing. In energetic terms, each salt bridge between an *e* and a *g* position contributes 1.5 KJ mol<sup>-1</sup> to coiled-coil stability.<sup>82</sup>

### 1.2.2 Attaining multivalency with the coiled-coil motif

Packing efficiency at the hydrophobic core of coiled coil-based peptides, together with the directionality of the *a* and *d* positions within the core, greatly influences the stability of the superhelix and the oligomerization state. It has been experimentally shown by Harbury's studies that the packing geometries of residues at *a* and *d* positions, respectively defined as parallel and perpendicular, govern the oligomerization state of parallel coiled coils (Figure 6).<sup>74</sup>



**Figure 6.** Intermolecular interactions within the coiled-coil motif. Helical wheel diagrams of **A:** a parallel, dimeric coiled coil; **B:** an antiparallel, dimeric coiled coil; **C:** a parallel, trimeric

coiled coil; **D**: a parallel, tetrameric coiled coil. Curved arrows indicate salt bridges, crossed-arrows depict hydrophobic interactions. **E**: Knobs-into-holes configuration in parallel dimeric, trimeric and tetrameric coiled coils; **F**: Knobs-into-holes configuration in antiparallel dimeric, trimeric and tetrameric coiled coils. Adapted from Apostolovic *et al.* with permission of Royal Society of Chemistry, copyright 2010.<sup>83</sup>

As mentioned in paragraph 1.2.1, in dimeric coiled coils isoleucine and leucine are preferred at *a* and *d* positions, respectively, whereas the opposite distribution favors the tetrameric assembly. In the case of trimers, instead, no discrimination has been observed, and they can be generated by  $a=d=Leu/Ile$ . A clear overview of the contributions of the amino acids at these two positions is given by Apostolovic *et al.* and is reported in Table 2.<sup>83</sup>

Degree of oligomerization	Orientation	
Amino acid sequence	Oligomerization number	Parallel <sup>a</sup> / Antiparallel
<ul style="list-style-type: none"> <li>Ile (<i>a</i>) and Leu (<i>d</i>)</li> <li>Substitution of four residues in the core of a potential tetramer with Ala creates cavities, which are destabilizing for tetramers and leads to dimer formation</li> <li>By inserting polar residues in the core</li> <li>Avoid Val and Asn at all <i>a</i> positions (induce oligomerization state &gt;2)</li> <li>Avoid hydrophobic residues at <i>e</i>, <i>g</i>, <i>b</i> and <i>c</i> position (induce oligomerization state &gt; 2)</li> </ul>	2	<ul style="list-style-type: none"> <li>Preferred interactions are <math>a : d'</math>, <math>d : a'</math>, <math>g : g'</math> and <math>e : e'</math></li> <li><math>a : d'</math> layers are Ala : Ile and Leu : Arg and <math>e : e'</math>, <math>g : g'</math> are attractive interactions (<math>g_n : e_{n+1}'</math> repulsive)</li> <li>Asn in the core specifies heterodimer formation through interhelical hydrogen bond <math>a : d'</math></li> </ul>
<ul style="list-style-type: none"> <li>Ile (<i>a</i> and <i>d</i>)</li> <li>Leu (<i>a</i> and <i>d</i>)</li> <li>Leu (<i>a</i>) and Val (<i>d</i>)</li> </ul>	3	<ul style="list-style-type: none"> <li>Trp in the core promotes antiparallel orientation as the bulky side chains can pack better this way</li> </ul>
<ul style="list-style-type: none"> <li>Leu (<i>a</i>) and Ile (<i>d</i>)</li> <li>Leu (<i>a</i>) and Ala (<i>d</i>)</li> <li>By inserting charged residues (E, K) at <i>b</i> and <i>c</i> position</li> </ul>	4	<ul style="list-style-type: none"> <li><math>a</math> (Met <math>\approx</math> Ile &gt; Val &gt; Leu &gt; Ala) and <math>d</math> (Met &gt; Ile &gt; Glu)</li> <li>Leu (<i>a</i>) and Ala (<i>d</i>)</li> <li>The insertion of Ala in the core of parallel tetramers promotes antiparallel assembly due to the formation of cavities</li> </ul>
<ul style="list-style-type: none"> <li>Trp (<i>a</i> and <i>d</i>)</li> <li>By inserting hydrophobic residues at all positions except <i>f</i> (the only position which is exposed while all others are buried)</li> </ul>	5	<ul style="list-style-type: none"> <li><math>a</math> (Leu &gt; Ile) and <math>d</math> (Met &gt; Gln &gt; Thr &gt; Val)</li> </ul>

<sup>a</sup> Note that > means that some amino acids are more preferred at *a* or *d* positions (better stabilizing tendency), while  $\approx$  means equally preferred (equal stabilizing tendency).

**Table 2.** Basic design principles to control the oligomerization state and helix orientation in *de novo* coiled coils. Adapted from Apostolovic *et al.* with permission of Royal Society of Chemistry, copyright 2010.<sup>83</sup>

Residues at *e* and *g* positions, despite being mainly responsible for the specificity of folding and the formation of homo- or heterooligomers, seem to also play a role in intrahelical interactions, due in part to attractive Coulombic forces and also due to their contributions to the stability of the coiled-coil; the latter occurs by increasing the hydrophobicity of the coiled-coil interface as, for example, in the case of protonated glutamic acid.<sup>82</sup>

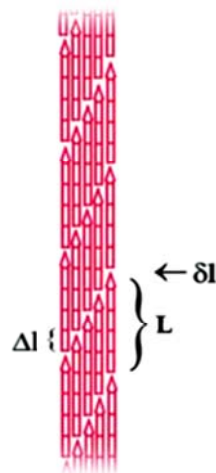
The number of helices is directly related to the stability of the structure. A higher number of helices translates into a broader hydrophobic interface, better shielding of the core from the aqueous environment, and an increase in the number of polar and ionic interactions via the amino acids at *b* and *c* positions. Interhelical interactions, however, do not only allow for a higher number of helices sharing the same hydrophobic core, but are also responsible for the longitudinal “growth” of coiled-coil oligomers into so-called  $\alpha$ -helical fibers.<sup>84</sup>

Fibrous structures are central to many biological processes and are of interest as bionanomaterials.<sup>85, 86</sup> The production of such structures is often available via self-assembly, a recurrent theme in biological systems. Molecular self-assembly allows the single, disordered building blocks of the system to form organized supramolecular structures as a consequence of specific, local interactions, without external intervention. It is a process mostly driven by noncovalent interactions and is observed in very small peptides and large proteins alike.<sup>87, 88</sup> Unlike all natural coiled coil-based structures, where the elementary units composing the fibers are mostly “blunt” ended, *de novo* designed  $\alpha$ -helical fibers are based on the “sticky” ends principle, a nomenclature adopted from nucleic acid research.<sup>89</sup> In the case of nucleic acids, digested DNA through the action of restriction enzymes presents one strand overhang to the other to form a very short single-stranded segment. This segment will easily repair with other ends of the same type, and thus are referred to as sticky ends. The same principle can be applied to *de novo* designed, coiled coil-based peptides: by constructing a peptide sequence which contains, at the C- or the N-terminus, an incomplete heptad repeat (for example: *defg(abcdefg)<sub>n</sub>a*), the sticky-end effect can be reproduced, and the peptide is able to form elongated fibers. It is likely, though, that the building block for these fibers is still the elementary oligomerization state of the designed coiled-coil (dimer or trimer). The Woolfson laboratory has widely exploited this concept by conceiving the first *de novo*-designed self-assembling heterodimeric fiber-forming peptide system (SAF-p1 and SAF-p2)<sup>90</sup> and developing it towards the formation of custom-designed supramolecular structures.<sup>91</sup> Potekhin *et al.* have proposed a structural model for the rational design of  $\alpha$ -helical fibers, starting from the formation of a five-stranded coiled-coil.<sup>92</sup> According to this model, a coiled-

coil peptide can assemble into an  $n$ -stranded ( $n$ ) fiber if the axial stagger between adjacent helices ( $\Delta l$ ) is equal to an integral number of heptads, according to the formula:

$$(L + \delta l) / n = \Delta l$$

where  $n$ =degree of oligomerization of the helical bundles, (typically 2, 3, 4, or 5),  $L$ =amino acid residues comprising the oligopeptide (generally 25–50),  $\delta l$ =residue equivalents (usually 0 or 1, for a single space between successive helices),  $\Delta l$ =axial stagger in residues between adjacent helices in the bundle. A graphical representation of this formula can be found in Figure 7.



**Figure 7.** Representation of the self-assembly of an  $\alpha$ -helical protomer into a five-stranded fibril according to the Potekhin model.

As mathematically elucidated above, several critical structural features must be taken into account when aiming to design a peptide sequence to self-assemble into  $\alpha$ -helical fibers: the coiled coil oligomerization state, the orientation between the adjacent helices (parallel/antiparallel) and the potential for lateral association. With these design principles in mind, several *de novo* coiled coil-based, fiber-forming peptides have been synthesized and have found applications in several fields.<sup>93, 94</sup>

### 1.2.3 $\alpha$ -Helical-based scaffolds for biomedical applications

Peptides based on the  $\alpha$ -helical secondary structure are relatively rigid scaffolds for the presentation of ligands to biological targets. The stiffness of the structure tightens their use to a

target-oriented ligand display, requiring specific positioning of the presented molecule on the backbone, with a considerable loss of entropy during target-ligand binding.<sup>95</sup>

Due to their high propensity to be found in helical conformation, two of the most used amino acid residues for this type of scaffold are proline and alanine. Polyproline chains (PP) adopt a well-defined left-handed helix of type II (PPII) and contain three residues per turn, aligning every third ring on the same face of the helix with a pitch of approximately 10 Å.<sup>96</sup> PP peptides have been used as a precision tool for the presentation of ligands in several applications. Examples are offered by the work of Yannick *et al.*<sup>97</sup> and Unverzagt *et al.*<sup>98</sup> The former has systematically studied the role of cationic and hydrophobic moieties in cell penetration using a PP backbone, demonstrating dramatic increases in uptake when up to six guanidinium groups were positioned on the polyproline helix, whereas only modest increases in cellular uptake were observed with the amine-containing polyproline compounds, showing that amphiphilicity plays a key role in enhanced cell translocation. The latter working group synthesized a competitive inhibitor against influenza virus by presenting multiple sialyl-*N*-acetylglucosamine side chains at predefined distances along a PP peptide. Most of the multivalent sialoglycopeptides exhibit increased binding proportional to the spacing of the ligands, when compared to monovalent compounds.

Regarding alanine, its high abundance in nature and its structural simplicity make it an ideal model building block for the study of conformational changes and aggregation behavior.<sup>99</sup> Alanine-rich peptides have been investigated, for instance, by Asakura *et al.* as scaffold candidates for the delivery of glutamic acid residues for bone-repair applications.<sup>100</sup> These peptides show enhanced cell binding through focal adhesions and enhanced expression of osteoblast-like cells. Liu and Kiick have used alanine-rich peptides as carrier molecules for carbohydrate presentation against cholera toxin.<sup>101</sup> Enhanced binding to the toxin was explained by the multivalent effect, which increases the local concentration of the saccharide ligand in the vicinity of the receptor.

The robustness of the coiled-coil motif has been experimentally explored, among others, by Falenski *et al.*, by loading up to twelve glycan moieties on a homo-dimeric coiled-coil model peptide.<sup>102</sup> Remarkably, the high degree of O-glycosylation was easily tolerated by this structural motif, and thermal stability analysis demonstrated the structural strength of the coiled-coil as a scaffold for the presentation of carbohydrate ligands. A great amount of work has been dedicated to elucidating the *de novo* design and oligomerization of the coiled-coil motif, but its use as a delivery system for biomedical applications has started to be taken into consideration only very recently.



## **1.3 CARBOHYDRATE-PROTEIN INTERACTIONS IN CELLS AND MEDICINE**

Most essential cellular processes, such as growth, differentiation, motility, morphology and communication, are mediated by extracellular signals received by and exchanged between the surfaces of cells. Some of these stimuli are received in the form of extracellular fluids, as in the case of hormones and neurotransmitters, but other signals are encoded in neighboring cell surfaces and exert their effect through direct cell-cell contact.

Cell surface carbohydrates represent information-rich binding sites for cellular intermediates and other cells and thus act as molecular recognition partners. Since most cell-cell and cell-environment communication occurs through carbohydrates, much current research is devoted to understanding their interactions with proteins and cellular compartments. For biomedical purposes, carbohydrates are often presented to their binding partners via synthetic scaffolds, to optimize interactions and fine-tune the signal response by means of the multivalent effect. A large portion of currently discovered multivalent interactions are mediated by carbohydrate–protein interactions, and their specificity is achieved by exploring the wide structural diversity of carbohydrates.<sup>103</sup> The next section attempts to describe the complex implications of carbohydrate-protein interactions and exploitation of these for research and medicine.

### **1.3.1 Biosynthesis of glycoproteins**

The attachment of sugar moieties during or after protein translation is complex.<sup>104</sup> It is a process that goes beyond the genome and is regulated by several factors via many elaborate pathways, to finally result in a mature glycoprotein that is either secreted by cells or becomes part of the membrane, cytoplasm or nucleus. The key event in the biogenesis of glycoproteins is clearly the formation of the glycopeptide bond, which determines the nature of the carbohydrate units and the protein's biological function. There are 13 different types of carbohydrate–amino acid combinations, arranged in five distinct groups, as shown in Figure 8. The most common carbohydrate–amino acid bonds in eukaryotes are:<sup>105</sup>

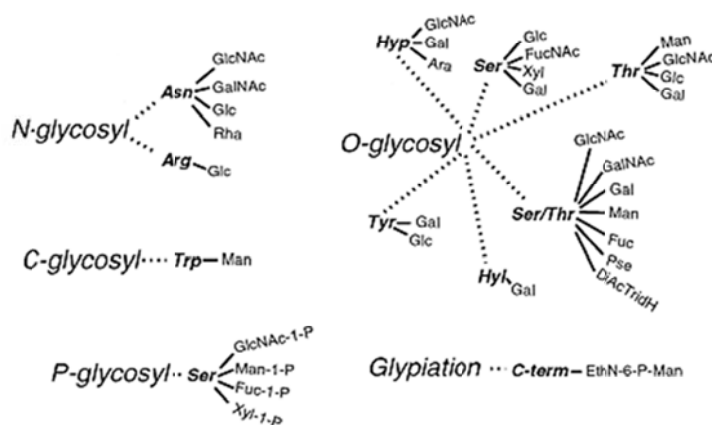
1. The O-glycosidic bond, between the hydroxyl group of the amino acid side chain (serine, threonine and tyrosine among the proteinogenic residues, and hydroxylysine and hydroxyproline among the others) and the terminal sugar.
2. The N-glycosidic bond, between the side chain amide of the amino acid asparagine and a glycan (exception: arginine with glucose).

- The bond via ethanolamine phosphate between the C-terminal amino acid of a protein and an oligosaccharide of the glycosylphosphatidylinositol anchor.

O- and N-glycosylation share some similarities but conform to very distinct rules. O-glycans do not share a common core, since O-glycosylation takes place exclusively post-translationally, but rather they have many different known core structures. The first monosaccharide attached in the synthesis of O-glycans is always an N-acetyl-galactosamine, but several different pathways are possible. N-glycosylation, on the other hand, only occurs on the asparagine residue within the sequence Asn-X-Ser/Thr, where X can be any amino acid except proline. It always starts with co-translational transfer of the dolichol diphosphate (PP-Dol)-linked oligosaccharide  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$  to the asparagine side chain; sugar molecules are attached via pyrophosphate linkages and the oligosaccharide chain is extended stepwise. In contrast to N-glycosylation, in which a complex glycan tree is coupled to the protein, O-glycosylation can result from the addition of even a single monosaccharide to the amino acid side chain.

Many eukaryotic proteins are anchored to the cell membrane, forming the glycocalyx, the outermost component of the cell surface. The glycosylphosphatidylinositol (GPI) anchor is a tail of glycolipids post-translationally attached to the C-terminus of various proteins.

CARBOHYDRATE NOMENCLATURE	
Abbreviation	Name
Ara	Arabinose
DiAcTridH	Diacetamido-trideoxyhexose
EthN-6-P-Man	Mannose-6-Phospho-Ethanolamine
Fuc	Fucose
Fuc-1-P	Fucose-1-Phosphate
FucNAc	N-Acetyl-Fucosamine
Gal	Galactose
GalNAc	N-Acetyl-galactosamine
Glc	Glucose
GlcNAc	N-Acetyl-glucosamine
GlcNAc-1-P	Glucose-1-Phosphate
Man	Mannose
Man-1-P	Mannose-1-Phosphate
Pse	Pseudaminic acid
Rha	Rhamnose
Xyl	Xylose
Xyl-1-P	Xylose-1-Phosphate



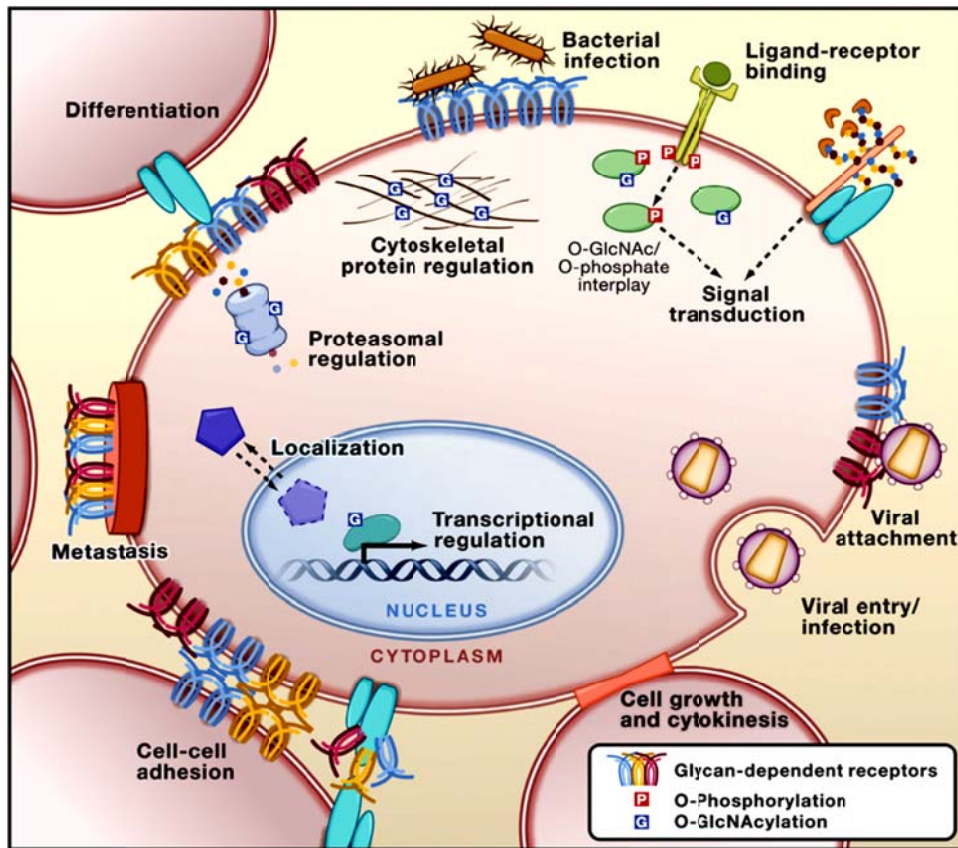
**Figure 8.** Diagrammatic representation of the most common types of sugar-peptide bonds and table summarizing the abbreviations shown in the sketches. Adapted from Spiro, with permission of Oxford University Press, copyright 2002.<sup>106</sup>

Protein glycosylation is controlled by rates of polypeptide translation and protein folding, localization of glycosyltransferases (enzymes responsible for glycopeptide bond formation), cellular sugar concentration and membrane trafficking. Therefore, single glycosylation sites on the same polypeptide can contain different glycan structures reflecting both the type and state of the cell in which they are synthesized.

Whether it is a single sugar or a complex glycan tree, glycosylation is an effective way of generating diversity in proteins and modulating their properties due to the inherent structural variations of glycans. For example, the bulkiness and the hydrophilic nature of carbohydrates often increase protein solubility and stability against proteolysis.<sup>107</sup> Sugars can stabilize the folded state of proteins, and this generally leads to higher melting temperatures. These effects typically depend highly upon the precise location of the sugar on the protein, so that different glycosylation sites influence protein stability in different ways. Due to steric hindrance, glycosylation restricts the range of conformations available to the neighboring peptides within the same protein, even when a defined change in secondary structure does not occur. For this reason, it generally contributes to protein folding by limiting the conformational degrees of freedom of the backbone around the site.<sup>108</sup>

### **1.3.2 Carbohydrate functions in cells**

Cells run on carbohydrates and glycoproteins are one of the most abundant and structurally diverse classes of molecules in nature. Carbohydrates are directly involved in almost all biological aspects of cell life and play an important role in almost every human disease. The functions of complex glycans are generally carried out at the multicellular level, as demonstrated by the fact that cells can grow in culture without them, but other simple sugars, such as the O-GlcNAc, are essential for mammalian cell homeostasis.<sup>109</sup> Glycoconjugates provide high structural diversity to proteins at the phenotype level, for various metabolic pathways, and for cell development, and glycoproteins are involved in both intracellular and intercellular communication, essential for tissue formation and growth. A synthesis of the processes in which carbohydrates are involved is offered by Figure 9.



**Figure 9.** Complex glycans at the cell surface are targets of microbes and viruses, regulate cell adhesion and development, influence metastasis of cancer cells, and regulate myriad receptor:ligand interactions. Glycans within the secretory pathway regulate protein quality control, turnover, and trafficking of molecules to organelles. Nucleocytoplasmic O-linked N-acetylglucosamine (O-GlcNAc) has extensive crosstalk with phosphorylation to regulate signaling, cytoskeletal functions, and gene expression in response to nutrients and stress. Adapted from Hart and Copeland, with the permission of Elsevier Inc., copyright 2010.<sup>110</sup>

As shown in Figure 9, surface glycans are often associated with cell adhesion and metastasis. In particular, there are alterations in the pattern of the glycocalyx that contribute to tumor proliferation, either directly, acting as signaling variations, or, more frequently, via modulating the function of the receptors they are attached to.<sup>111</sup> An example is offered by the action of certain glycosyltransferases which regulate the Notch's pathway by its ligands, affecting many developmental events.<sup>112</sup>

Glycans are also greatly involved in immunity, infection and inflammation:<sup>113</sup> nearly all microbes and viruses infect the human body by attaching to a specific cell-surface glycan or interacting with one of the many surface lectins (carbohydrate-binding proteins) of the host

cell. One of the main reasons that facilitate this event is molecular mimicry: surface carbohydrates of many pathogens have a very similar chemistry to the mammalian components of the eukaryotic glycocalyx, and this similarity often allows bacteria and viruses to evade immune surveillance, like in the case of the sialic-acid capsule of group B *Neisseria meningitidis* and neural cell adhesion molecules.<sup>114</sup> The biochemical tolerance and alteration of the host glycosylation pattern during disease/infection are both components that must be considered to better understand the sugar-mediated cell invasion. For these reasons, carbohydrates are not perceived by the body as highly immunogenic and their use as a sole element for a vaccine preparation does not guarantee the activation of the immune response. This is why carbohydrate-based vaccines frequently contain one or more immune-stimulatory compounds, such as peptide epitopes, providing appropriate stimuli to the host's immune system. In this way, the cell-mediated immunity would provide immunological memory, while the carbohydrate antigen would allow for the production of specific antibodies.

### **1.3.3 Multivalency in carbohydrate-protein interactions**

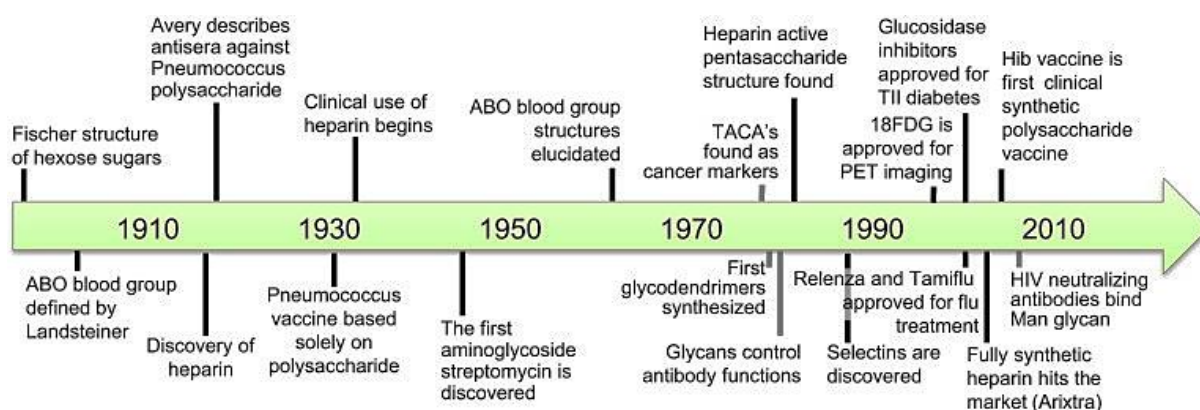
Whenever the topic is carbohydrate and cells, it is also important to take into consideration the nature of the carbohydrate-protein interactions: the two components are linked by a number of relatively weak interactions that ensure specificity but permit unlinking when needed. The best examples are offered by the carbohydrate-binding protein family of lectins, whose main role is to facilitate cell-cell communication. Due to the weak nature of the interaction with carbohydrates, a lectin is usually formed by more than one binding site which often interact with an array of glycans expressed on the membrane of another cell. The essence of this binding modality resembles the action of Velcro:<sup>115</sup> each bond is relatively weak but the complex of all interactions is strong. Binding affinity and degree of carbohydrate substitution and branching have been widely analyzed and frequently correlated.<sup>116</sup> A good example is offered by a triantennary inhibitor of the lectin asialoglycoprotein receptor (ASGPR), which, with its three GalNAc moieties, shows higher inhibitory potency than the biantennary counterpart.<sup>117</sup>

The phenomenon of multivalency and formation of multimeric architectures containing several recognition sites is very frequent in nature, since it allows for the compensation of weak individual noncovalent interactions. The proposed model of binding for mammalian lectins is a mechanism of multivalent recognition, in which avidity can be further increased by matching appropriate conformational arrangement of the binding site in the lectin oligomer

with the spacing of carbohydrate ligand.<sup>118</sup> Another example is offered by the mannose-binding protein-specific antibody 2G12, which achieves its recognition via multivalent interactions, rather than recognition by a single high affinity site.<sup>119, 120</sup> In conclusion, protein-carbohydrate recognition at molecular levels is always enabled by a synergistic combination of several cooperative binding events.

### 1.3.4 Use of carbohydrates in medicine

Over the past two decades, there have been many improvements in tools and techniques for the study and synthesis of carbohydrates and their use as therapeutics.<sup>121</sup> Recognition of the importance of glycans in the medical community dates back to the discovery of the three blood types (A, B and 0), in 1900.<sup>122</sup> The identification of the chemical differences and identities of the blood type components had to wait other 50 years for the work of Kabat, Leskowitz, Morgan and Watkins, which demonstrated that the main constituent of the H antigen was the monosaccharide fucose, to which either N-acetylgalactosamine (GalNAc) or galactose (Gal) were added to A and B antigens, respectively.<sup>123-125</sup> Since this very early discovery, carbohydrates made their official debut into various fields of medicine and pharmacy, such as angiology, immunology and diabetology.



**Figure 10.** Timeline of glycans in medicine. Adapted from Hudak and Bertozzi, with permission of Cell Press, copyright 2014.<sup>126</sup>

As shown in Figure 10, carbohydrates were introduced into the field of immunology with the identification of the polysaccharide-based coating from the *Pneumococcus* capsid that could

react with type-specific antisera from patients infected with the pathogen.<sup>127</sup> This discovery opened the gates for the new concept that carbohydrates can be used as a main component for vaccine development.<sup>128, 129</sup> In particular, their use in the vaccine Pneumovax (PPV23), containing over twenty capsular polysaccharides from *Streptococcus pneumoniae*,<sup>130</sup> demonstrates that glycans are indeed suitable to efficiently stimulate and “train” the antibody response for effective vaccination. However, the low molecular weight and the inherently T-cell-independent nature of carbohydrates make them poorly immunogenic.<sup>131</sup> The first response to this limitation was offered by conjugating carbohydrates to carrier molecules, but this approach also failed to induce a sufficient T-cell response in human patients.<sup>132</sup> A solution was presented by adding a promiscuous T-cell epitope to the vaccine components to trigger, simultaneously, the T-cell-dependent immune response and specific antibody production via the glycopeptide antigen.<sup>133</sup> This successful design brought about the development of vaccines against, for example, *Neisseria meningitidis* (Menactra), *Streptococcus pneumoniae* (Pneumovax), *Haemophilus influenzae* type b (Hib; Hiberix, Comvax), and *Salmonella typhi* (TYPHIM Vi).<sup>134, 135</sup>

The chosen carrier system also plays an important role in carbohydrate-based vaccine production. The above-mentioned vaccines utilize a protein as carrier, generally a deactivated version of a known toxin. Recently, greater attention has been paid to fully synthetic vaccines. The most successful example may be offered by the production of the Cuban Hib vaccine, the first clinically approved carbohydrate-based vaccine, the synthesis of which was described by Verez-Bencome *et al.* and is based on conjugating an antigenic tetrasaccharide from Hib to the tetanus toxin.<sup>136</sup> Equally impressive is the work by the Seeberger group to generate a synthetic, glycan-based anti-malarial vaccine, which reduces many sources of tissue damage normally found during infection.<sup>137, 138</sup>

Beyond their use in immunology, the interaction of carbohydrates with their main binding partners, the lectins, is one of the most relevant phenomena that must be taken into account for therapeutic purposes. The discovery of selectin in the 1980s represented a major step forward in the field of glycobiology and inspired decades of therapeutic development. It must be remembered that protein-carbohydrate interactions are very weak in nature, and therefore monovalent sugars as targets for lectin binding are not suitable to be considered as drug candidates. For this reason, multivalent glycoconjugates have been developed taking advantage of the “cluster glycoside effect” to strengthen lectin avidity,<sup>139, 140</sup> resulting in glycopolymers, glycodendrimers and glyconanoparticles.<sup>141, 142</sup> The clinically relevant dendrimeric scaffold STARFISH<sup>143</sup> is an oligovalent dendron-like compound bearing a high

number of selected glycans. The same STARFISH dendrimer has been used to inhibit or image specific carbohydrate-binding proteins. For example, Rele *et al.* demonstrated that a high valency polyethylene oxide (PEO) star dendrimer bearing highly sulfated lactose ligands can selectively recognize selectin and reduce inflammation when administered in a mouse model.<sup>144</sup> Also, the architecture of the multivalent structure makes a major contribution to the activity of the final glycoconjugate, as shown, for instance, by the Kiessling research group, which demonstrated that a designed multivalent scaffold was essential to enhance avidity and specificity by varying scaffold shape, size, valency, and the density of the binding elements.<sup>145</sup> A further example is offered by Scheibe and Seitz, who provided multivalent PNA-DNA complexes bearing sugar conjugates presented in a spatially-defined fashion.<sup>146</sup> the periodicity and flexibility of the nucleic acid helix allowed for the precise determination of number and spatial alignment of the presented carbohydrate ligands. It is now generally understood that combining the specificity imparted by the carbohydrates and the multivalent effect offered by a synthetic scaffold is a promising strategy for optimizing the therapeutic effects of carbohydrate-based drugs.



# 2

**Synthetic Glycopeptides as Multivalent Scaffolds for Carbohydrates:**  
from receptor targeting to vaccines exploiting sugar-protein interactions

## APPLIED METHODS

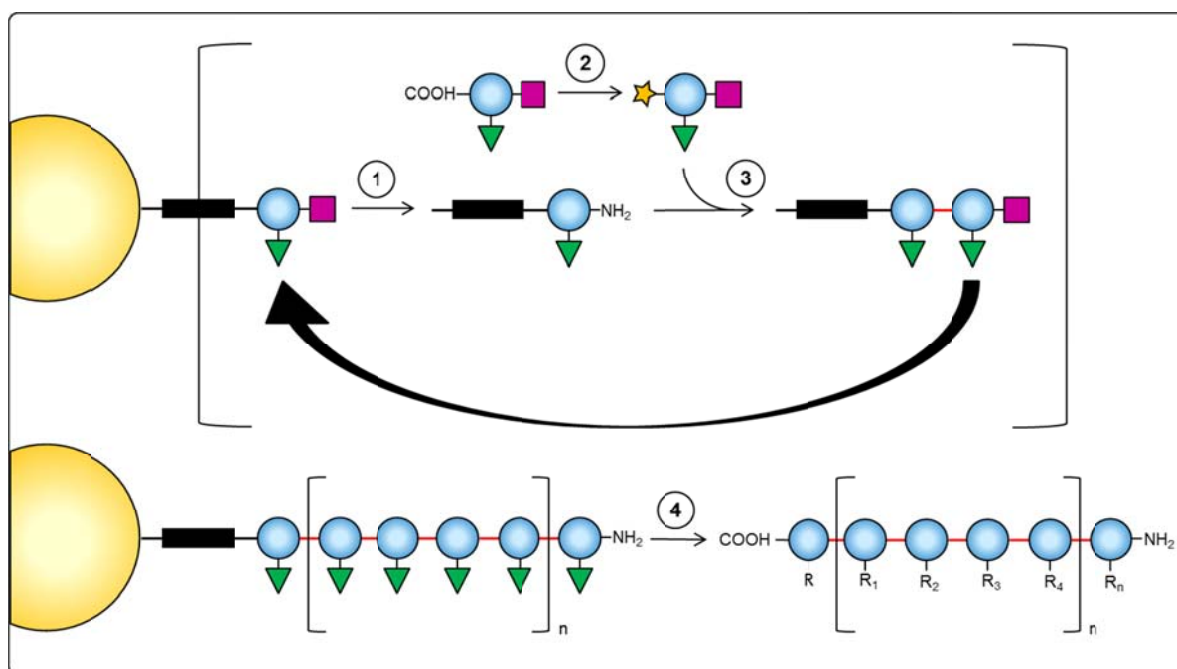
- 2.1 Peptide production and purification
- 2.2 Further analytical methods
- 2.3 Biological methods and techniques

## 2.1 PEPTIDE SYNTHESIS AND PURIFICATION

### 2.1.1 Solid-phase peptide synthesis

Peptides are synthesized via a series of couplings of the carboxylic group of an amino acid to the amino group of another. The chemical synthesis of peptides is conducted from the C-terminus to the N-terminus and, to avoid side reactions, the use of protecting groups on the amino acid side chains is generally required.

Excluding the biological methods of peptide production, such as extraction or bacteria/yeast expression, there are two main chemical approaches, the classical liquid-phase synthesis and the solid-phase synthesis, and two main strategies, tert-butyloxycarbonyl (Boc) or fluorenylmethyloxycarbonyl (Fmoc) based. All peptides generated for this thesis were synthesized by solid-phase peptide synthesis (SPPS) with an Fmoc-based strategy, also called Merrifield synthesis.<sup>147</sup> The main principles are described below.



**Figure 11.** Scheme of the SPPS. Yellow sphere: resin bead; black rectangle: resin-peptide linker; blue circle: amino acid; green triangle: side chain protecting group; purple square: Fmoc group; yellow star: carboxylic acid activation site; red line: peptide bond.

As depicted in Figure 11, Fmoc-based SPPS is a repetition of three main steps: 1: Fmoc removal; 2: activation of the next amino acid; 3: coupling. The Fmoc group is cleaved under basic conditions (typically 20% - 40% piperidine) while the protecting groups of the side

chains are not base labile. This orthogonal strategy allows for the exposure of the amino group of the amino acid anchored on the resin, while avoiding any unwanted reaction with the side chain. The activation of the carboxylic acid of the following amino acid is performed in a separate tube with specific activating agents, according to the characteristics of the amino acid in question, such as bulkiness and solubility. The most commonly used combination is diisopropylcarbodiimide (DIC) and hydroxybenzotriazole (HOBt). The activated amino acid is added to the reactor for a certain time that can vary from 30 minutes to overnight, according mainly to the accessibility of the functionalities in solution. When all the amino acids are coupled together, strong acidic conditions, generally 95% trifluoroacetic acid (TFA), can cleave the peptide from the resin and remove the side chain protecting groups, simultaneously. After solvent removal, the peptide is ready to be purified.

### **2.1.2 Reversed-phase high-performance liquid chromatography**

The most used technique for synthetic peptide purification is high-performance liquid chromatography (HPLC). It is a chromatographic technique for separation of components from a mixture, their quantification and isolation. The reversed phase (RP)-HPLC, allows for separation of compounds according to their polarity. The most hydrophilic compounds elute first, followed by compounds with decreasing polarity. This is due to the reduced interactions of hydrophilic compounds with the hydrophobic stationary phase of the column when the polarity of the mobile phase is decreased (e.g., increased percentage of acetonitrile over water). Elution is monitored by UV-detection and the eluents can be collected and analyzed separately.

### **2.1.3 Electrospray ionization-mass spectrometry**

Mass spectrometry (MS) is an analytical technique to identify the type of molecule analyzed by measuring the mass-to-charge ratio of gas-phase ions. Electrospray ionization (ESI) produces ions by applying high voltage to a liquid to create an aerosol. This type of MS is particularly useful when analyzing proteins and peptides, since it overcomes their propensity to fragment during ionization. The typical solvent mixture for ESI-MS is water with a volatile organic compound (very often acetonitrile or methanol), because ion formation requires extensive solvent evaporation. Evaporation of the solvent results in the formation of charged

droplets of increasingly reduced size. In the case of peptides, this technique may produce multiply charged ions, extending the mass range to accommodate the kDa-MDa orders.

## 2.2 FURTHER ANALYTICAL METHODS

### 2.2.1 Circular dichroism

Circular dichroism spectroscopy is a widely-used method to gain low-resolution structural information about proteins and peptides in solution.<sup>148</sup> In particular, it is frequently used to determine protein and peptide conformations and conformational changes, and it has been employed to characterize the secondary structure of all peptides produced in this study.

Circularly-polarized light consists of an electric and a magnetic field, perpendicularly oscillating. The electric field does not change its strength but only its direction, in a rotary fashion, so that the tip of its vector describes a helix along the direction of propagation. Circular dichroism (CD) results from the interaction of a chiral chromophore with circularly polarized light and measures the difference between the absorption of left-handed circularly polarized light (Lcp) versus right-handed circularly polarized (Rcp) light. The absence of a difference yields an absence of CD signal, which is indicative of the absence of a regular structure; any difference, positive or negative, results in variation of the absorption of the light and a consequent positive or negative intensity signal in the CD spectrum. When circularly polarized light hits an optically active compound in solution, for example a peptide with a definite structure, various parameters will vary between the right and left polarized light: the speed of polarization, their wavelength and the extent to which they are absorbed. Particularly, this last parameter is the main determinant that results in a CD signal, according to the formula:

$$\Delta A = A_L - A_R = \epsilon_L Cl - \epsilon_R Cl = \Delta \epsilon Cl$$

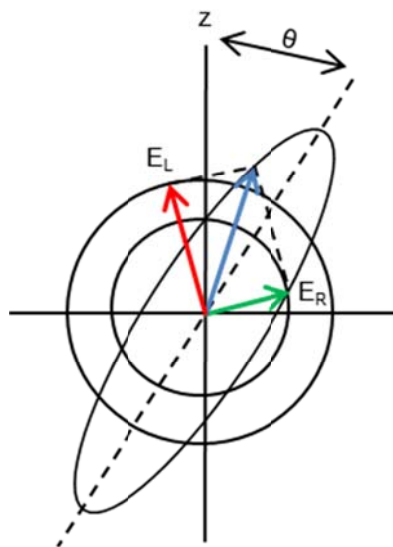
where the difference in the absorption of the left (L)- and right (R)-polarized light ( $\Delta A$ ) is given by the difference between the decadic molar extinction coefficient of the solute for L and R (respectively  $\epsilon_L$  and  $\epsilon_R$ ), each one multiplied by the molar concentration of the analyte (C) and by the optical path length (l). Therefore, according to the formula,  $\Delta A = \Delta \epsilon Cl$ .

This difference explains why the electric vectors of circularly polarized light outline an ellipse, rather than oscillating within the plane. When the electric vectors of the L and R components have the same direction, the main axis of the ellipse is the sum of their magnitude; when their directions are opposite, their difference results in the minor axis of the ellipse. The ratio of major and minor axes is defined as ellipticity.

When the L and R circularly polarized components are in-phase with equal amplitude and wavelength, the superimposition of their vectors results in a line. This is what happens before the beam passes through the sample in a CD spectrometer. When the beam is perturbed by an asymmetric sample, the two circularly polarized components (L and R) will be absorbed differently, and the projection of their vector will describe an ellipse, rather than a line. The rotation angle  $\theta$  is defined as:

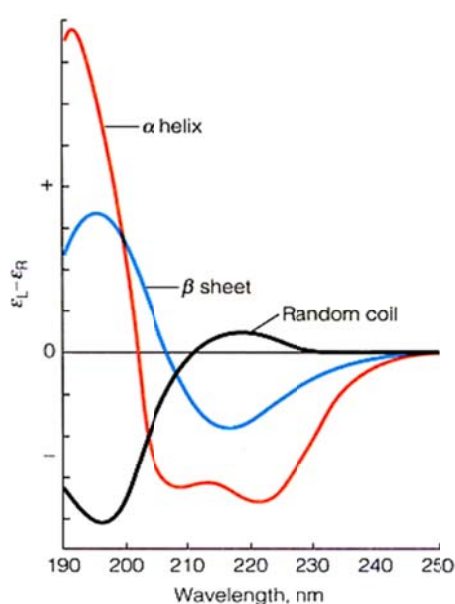
$$\theta = \frac{\pi C}{\lambda} (\epsilon_L - \epsilon_R)$$

The projection of the field vector defining the ellipse can be found in Figure 12.



**Figure 12.** Ellipse described by the left and right polarized components when perturbed by an asymmetric chromophore. Red arrow ( $E_L$ ): field vector of left polarized light. Green arrow ( $E_R$ ): field vector of right polarized light. Blue arrow: projection of the superimposed vector, describing the ellipse.  $\theta$ : degree of optical rotation.

Peptides and proteins are ideal molecules to be studied by circular dichroism, because they consist of chiral amino acids connected via peptide bonds, which serve as chromophores and interact with the beam of light. The peptide bond, or amide bond, most efficiently absorbs light at wavelengths shorter than 250 nm. To gain information about the secondary structure of a certain peptide, the result of CD spectroscopy, the CD spectrum, must be considered to be a series of bands corresponding to different electronic transition states, from the ground state to an electronically excited one. In peptides and proteins, the two characteristic transitions are the  $\pi$ - $\pi^*$  transition, occurring at 190 nm, and the weaker  $n$ - $\pi^*$  transition, at 210 nm. The singularity of the hydrogen-bond network characteristic of each secondary structure results in different CD spectra, as shown in Figure 13.



**Figure 13.** Characteristic CD spectra of  $\alpha$ -helical (red),  $\beta$ -sheet (blue) and random coil (black) conformations.

The CD spectra of a certain peptide or protein can be deconvoluted as the sum of the  $\alpha$ -helix,  $\beta$ -sheet and random coil contributions:

$$CD_{\text{protein}} = AX_{\alpha\text{-helix}} + BX_{\beta\text{-sheet}} + CX_{\text{random coil}}$$

where X represents the contribution of each secondary structure to the global spectrum, and A, B and C are the data points corresponding to the ideal spectra of the three conformations. Strong minima at 222 nm and at 208 nm, together with an intense maximum at around 192-195 nm, are representative of an  $\alpha$ -helix, since it correlates with the hydrogen-bonding

environment of this conformation; the  $\beta$ -sheet conformation is represented by a single minimum between 210 nm and 220 nm, and a positive peak of similar intensity between 190 nm and 200 nm. Many denatured proteins and peptides, oligopeptides and polypeptides possess a CD spectrum typical of a random coil, with a single minimum near 195 nm and very weak positive, broad bands between 220 nm and 230 nm. Such a CD profile means that their building blocks are still bonded together but there is no well-defined hydrogen-bonding network between them.

The interpretation of a CD spectrum is not always straight forward, although it is generally easier for short model peptides. In nature, proteins are generally composed of more than one type of secondary structure, and the corresponding CD profile is a proportional combination of their contributions.

### **2.2.2 Transmission electron microscopy**

Some of the peptides involved in this study have the tendency to form higher-order structures which differ from amorphous aggregates, because they show regularity and consistency in the interaction between their building blocks. To be able to observe these supramolecular architectures, the light microscope does not offer appropriate resolution. Thus, in this work, these peptides are analyzed via transmission electron microscopy (TEM).

Electrons possess dualistic particle-wave properties and can therefore be used as a radiation source. This principle developed into the first prototype instrument called transmission electron microscope. Electron microscopy is a technique which allows for the observation of samples at high resolving power. Compared to more standard visible light, electron beams possess up to a 100,000-fold shorter wavelength of a few picometers, offering high-definition images as result of the interaction between the transmitted electrons and the sample.<sup>149</sup>

To acquire an image via TEM, a high-voltage electron beam (100 kV) is focused via electromagnetic and electrostatic lenses and transmitted to the sample, which is pretreated (negative staining or cryo-TEM) to be partially permeable to electrons. The structural information carried by the electron beam are finally transmitted to an imaging system of the microscope, where the images can be magnified via electromagnetic lenses, recorded and developed onto TEM micrographs.

When an electron beam hits the sample, it can be elastically bounced without loss of energy (if it hits the nucleus of an atom), or it can interact by striking an electron, and transfer some

of its energy to the atom. In the first circumstance, the angle of the elastic bounce will be dictated by the law of conservation of momentum; in the second case, the residual energy transferred by the electron beam is instead random, and, when the beam reaches the imaging system, possesses unknown energy and angle, which leads to noise.

TEM imaging uses the elastic scattering to generate transmission contrast. Obtaining the right degree of contrast in TEM is not as straight forward as for the light microscope, since electrons interact with matter in a different way compared to light. Moreover, the electron-scattering of biological samples, deriving mostly from C, H, O and N, is weak; these samples must, therefore, be coated with heavy metal ions to increase their contrast, since they have higher weight and nucleus radius and can strongly scatter the electron beam.

The most established technique to create contrast in an organic specimen is negative staining.<sup>150</sup> The method consists in imaging the sample using an optically opaque fluid and stain the background, rather than the sample. In this work, the most used negative stain was phosphotungstic acid. It scatters the electrons very well and quickly absorbs to biological samples. Practically speaking, to prepare a peptide sample with negative staining TEM few  $\mu\text{L}$  of sample solution is added to a carbon support and allowed to absorb for a short time. The droplet is consequently removed by touching it with the edge of a blotting paper and a droplet of staining solution is pipetted on the carbon support containing the sample for few seconds. The grid is allowed to air-dry and can subsequently be analyzed.

Despite being the indicated technique for the analysis of peptides and proteins, negative staining TEM provides only information about the outer sample structure. To better study the core of the peptide aggregates, any contrast reagent must be avoided. For this reason, cryo-transmission electron microscopy (cryo-TEM) has been developed and it is now considered the best method to allow the observation of specimens that have not been stained or fixed in any way, showing them in their native environment. With this methodology, the sample is immobilized by freezing (as suggested by the name “cryo”). This allows to restrain the structural characteristics of the analyzed protein/peptide aggregate and to observe it in its native-like conformation.

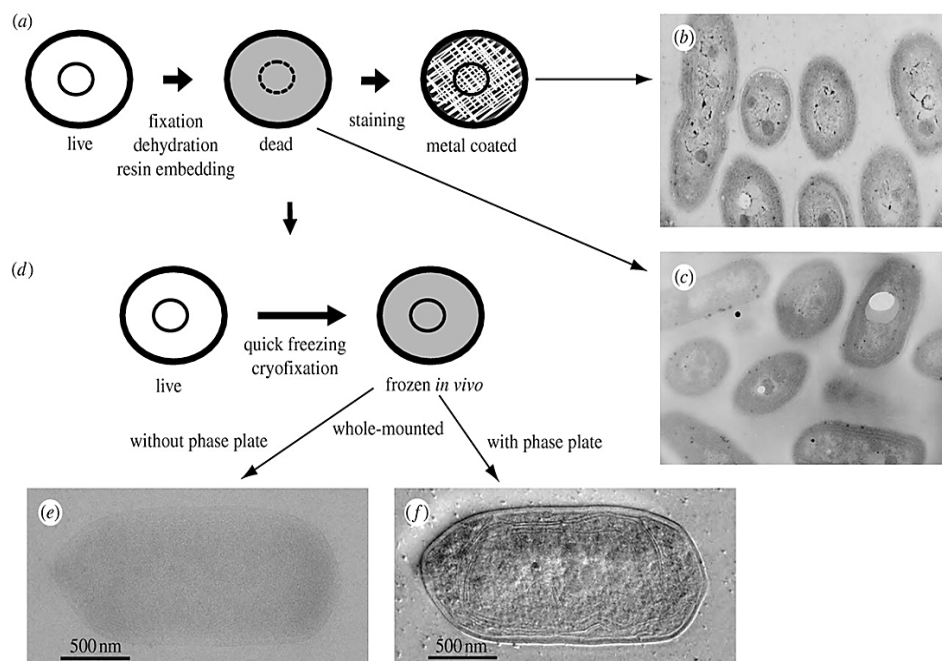
The freezing procedure must be quick and must allow for the electron beam to be absorbed. The faster the freezing, the higher the chance that also the water will be instantly solidified as a disordered, amorphous solid. This process is called vitrification. The poor ordering degree of the water molecules in the sample makes it transparent for the electron beam and therefore



available for TEM analysis. This means that the sole analyte responsible for the scattering of the electron beam is the vitrified protein/peptide aggregate, since no stain is included. As a consequence, the contrast is very low, due to the lack of heavy atoms. Nonetheless, cryo-TEM remains a very efficient technology to study the architecture of proteins and peptide aggregates at molecular resolution and in a native-like state.

The sample preparation for the cryo-TEM is more challenging than for the negative staining. The specimen is generally frozen in liquid ethane or nitrogen and must be prepared on a thin layer, to avoid further diminishing the contrast. Moreover, also the grid must be kept cold throughout the analysis, to avoid melting and evaporation.

The results obtained by negative staining TEM and cryo-TEM on a biological sample can be compared in Figure 14.



**Figure 14.** Comparison between negative staining TEM and cryo-TEM. A standard sample preparation and an improved sample preparation using a cryotechnique. TEM micrographs associated with the preparation techniques are compared (b,c,f). (b) Conventional TEM images of a standard cyanobacterial cell sample (100kV). (c) Conventional TEM images of a standard cyanobacterial cell sample without staining (100kV). (d) A sample preparation using quick freezing. (e) A conventional TEM image for a vitrified cyanobacterial cell with deep defocusing (300kV). (f) An HDC-TEM image for a vitrified cyanobacterial cell (300kV). Adapted from Nagayama and Danev, with permission of The Royal Society publishing, copyright 2008.<sup>151</sup>

## 2.3 BIOLOGICAL METHODS AND TECHNIQUES

### 2.3.1 Flow cytometry

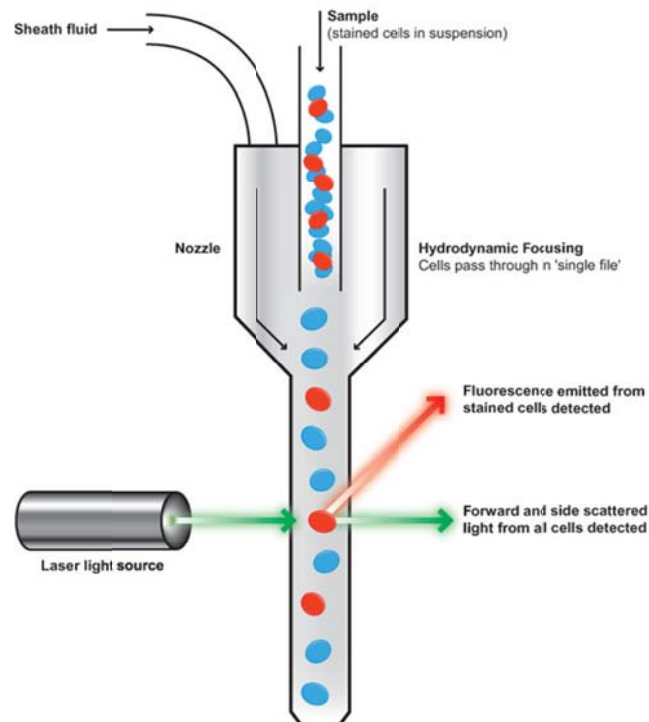
The interaction of some of the peptides employed in this thesis with cells has frequently been analyzed by means of flow cytometry. This biophysical technology allows for the fast measurement of multiple characteristics of individual cells as they flow in a fluid stream through a beam of light. This feature makes flow cytometry a very powerful tool for detailed analysis of complex cell populations in a short period of time.

Flow cytometry can be used to detect and analyze a particle's relative size, relative granularity or internal complexity, and relative fluorescence intensity. This last property is the most relevant for our purpose: the study of the interaction of labeled peptide with cells. The results are determined using an optical-to-electronic coupling system that records how the cell scatters incident laser light and emits fluorescence.<sup>152</sup>

A flow cytometer is composed of three systems: the fluidic system, which transports the cells to the laser beam for analysis; the optic system, which allows for laser illumination of the sample, while filters direct the light signals towards the detectors; and the electronic system, which converts light signals into electronic ones for data processing. When cells (or particles between 0.2  $\mu\text{m}$  and 150  $\mu\text{m}$ ) in solution pass through the light source, they scatter the light and, if an appropriate wavelength is applied, can emit fluorescence. Scatter and fluorescence signals are collected by means of a series of lenses and sent to appropriate detectors, which produce electronic signals proportional to the optical ones. Cells in suspension are drawn into a stream created by a surrounding sheath of isotonic fluid that creates laminar flow, allowing them to pass individually through an interrogation point. Thus, data are collected for each single cell/particle that passes through the light beam and the data sent to software for analysis. The two main data sets to be analyzed derive from light scattering and fluorescence emission. Light scattering occurs when the laser hits the cell and its extent is proportional to the cell's size and internal properties. In particular, forward-scattered light (FSC) is proportional to the cell surface area, while side-scattered light (SSC) is a function of the granularity of the cell. The combination of the information offered by FSC and SSC allows for cell type differentiation when a heterogeneous population is analyzed.

Fluorescence emission occurs when the light absorbed by a fluorophore causes an electron to be excited to a higher energy level; this electron then quickly decays to its ground state, emitting the excess energy as a photon of light. Fluorescence is of particular utility to study

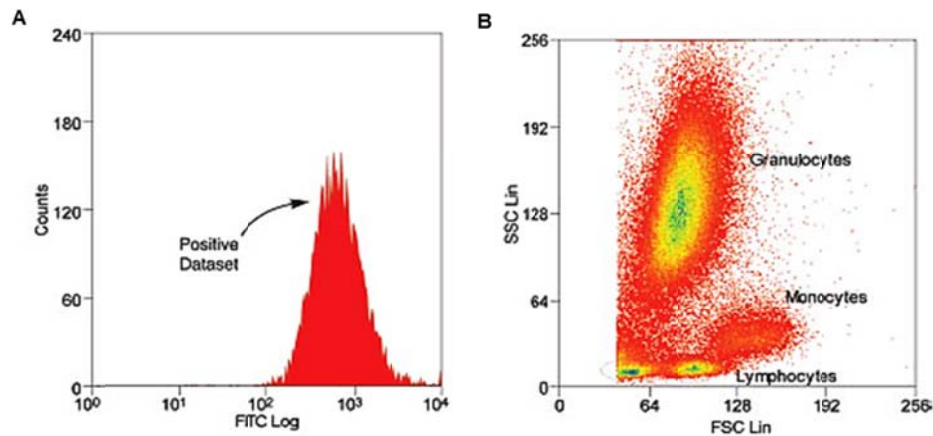
cell-protein/peptide interactions. Fluorescent dye-labeled peptides can be incubated with the selected cell type prior to flow cytometry analysis. The cells which “uptake” the fluorescent peptide (via binding or internalization) will consequently be fluorescent and can be counted and sorted. Comparison between the amount of fluorescent cells and non-fluorescent cells will give a quantitative estimate of the degree of interaction of a specific peptide with the cells. The basic working principles of flow cytometry are shown in Figure 15.



**Figure 15.** Main components and working principle of a flow cytometer. The sample is a mixture of stained and not-stained cells. A laser set with a specific wavelength delivers a light beam to the cells in the thin chamber. The scattered light and the emitted fluorescence are detected and the light signal gets converted to an electronic one for analysis. Adapted from Abacam, *Flow Cytometry Guide*.<sup>153</sup>

The data generated by flow cytometric analysis can be plotted in a one-parameter histogram, or in a two- or three-parameter dot plot. Examples can be found in Figure 16. The histogram displays either the relative fluorescence or the light scattering intensity on the x-axis and the number of events (cell counts) on the y-axis. It shows the number of cells that possess the previously selected marker of interest. These cells form the “positive dataset”.

In a two-parameter dot plot, as for example shown in Figure 16, the fluorescence and the light scattering properties of the cell population are shown on the two axes, respectively. A color map generally indicates the cell count.



**Figure 16.** Plotted flow cytometry data. A: one-parameter histogram showing the FITC log (fluorescence mean) on the x-axis. B: three-parameter data plot, showing fluorescence on the x-axis, light scattering intensity on the y-axis and color-coded cell count. Adapted from AbD Serotec’s guide: *Introduction to Flow Cytometry*, Chapter 1: Principles of flow cytometry.

On a density plot each dot represents an individual cell passed through the instrument. This implies that all cells, including amorphous cells and dead cells, will be taken into account. The advantage of this type of data plotting is that it displays also the light scattering intensity: FSC is closely related to cell size and, therefore, bigger cells will have stronger FSC signals; SSC is connected to the properties of the intracellular organelles and particles and, therefore, it is more sensitive to variations in membrane, cytoplasm and nucleus. The combination of these signals can be used to distinguish different cell populations (as in the case of Figure 16) or living cells from lysates.

Moreover, the different region of the intensity blot can also be separated according to the fluorescence intensity, obtaining subsets termed “gates”.

### 2.3.2 Confocal fluorescence microscopy

As explained above, flow cytometry offers the possibility to verify, for example, whether a labeled peptide interacts with a chosen cell type. The disadvantage of this technique is that it

does not differentiate between peptides that bind to the cell membrane and peptides that are instead internalized within the cytosol.

To overcome this limitation, it is necessary to use an imaging technique that allows the observation of cells in real time and the identification of the position of the peptides by colocalization with the fluorescence signal. This type of analysis can be performed by means of a confocal fluorescence microscopy.

Confocal microscopy is an imaging technique with increased optical resolution compared to widefield microscopy. The difference is due to the presence of a spatial pinhole positioned at the confocal plane of the lens, whose effect is to eliminate out-of-focus light and obtain a three-dimensional reconstruction of the analyzed specimen. The pinhole ensures that only the light produced by fluorescence very close to the focal plane will be detected, and improves optical resolution by enabling the collection of a series of optical sections of thick specimens. The light emitted by the laser (excitation source) passes through the pinhole opening at the confocal plane and through another pinhole in front of the detector. After the excitation light is reflected by a dichromatic mirror and scanned across the sample in a certain focal plane, the fluorescence emitted by the sample passes back to the mirror and is focused at the detector pinhole. Fluorescence occurring above and below the objective focal plane is not confocal with the pinhole, and, therefore, only a very limited amount of it is detected by the photomultiplier. Thus, the out-of-focus emitted fluorescence does not contribute to the resulting image.

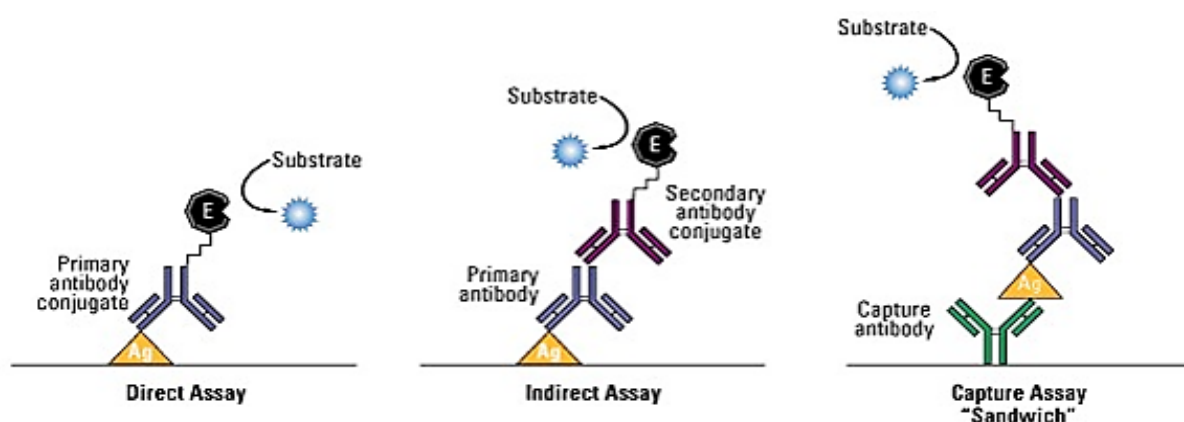
Besides much clearer images, a confocal microscope allows collection of optical sections in transverse planes. Vertical sections of the z-plane result in layers of images of the specimen at different distances from the optical axis and allows for a more detailed analysis of the intracellular distribution of the sample.

### **2.3.3 Enzyme-linked immunosorbent assay**

The enzyme-linked immunosorbent assay (ELISA) is a biochemical and immunological technique to detect and quantify the presence of an antigen, peptide, protein or antibody in a sample mixture. During this study, ELISA has been employed to analyze the binding of specific antibodies to a peptide epitope, displayed either monovalently or in a multivalent fashion. The general procedure of the ELISA and its relevance for the investigations conducted in this thesis is depicted below.

In an ELISA, an antigen is fixed on a solid support (specifically or non-specifically) and a detection antibody is applied to bind the antigen. This detection antibody can be covalently linked to an enzyme, or can itself be detected by a secondary antibody that is linked to an enzyme through bioconjugation. The enzyme catalyzes a reaction from a given substrate that normally generates a detectable signal (most commonly a change in color or fluorescence emission) that can be photometrically quantified. The intensity of the signal is proportional to the amount of detection antibody bound to the antigen. Between each step, the solid support, typically a 96-well polystyrene plate, is repeatedly washed to remove non-specifically bound material. This makes the ELISA a powerful tool for measuring an antibody's avidity to bind its substrate.

The methods of immobilization and the detection of the antigen are the two main variables within the different types of ELISAs. The immobilization of the antigen of interest can be accomplished by direct adsorption to the plate or by fixing on the plate a first antibody and then adding the antigen to be captured by it. This is the main difference existing between "standard" and "sandwich" ELISA. In the same fashion, the detection of the antigen can occur directly, via an enzyme-conjugated primary antibody, or indirectly, via an enzyme-conjugated secondary antibody which binds to the primary unlabeled antibody attached to the antigen. These two ways of proceeding are respectively referred to as "direct" and "indirect" ELISA.<sup>154</sup> Because multiple secondary antibodies can bind to each primary antibody, this approach allows multiple reporter molecules to localize to each antigen, thereby amplifying the signal and increasing the sensitivity of antigen detection. For this reasons, the indirect ELISA was preferred during the course of this research. A diagram of common ELISA formats is illustrated in Figure 17.



**Figure 17.** Common ELISA formats. In the assay, the antigen of interest is immobilized by direct adsorption to the assay plate (direct ELISA) or by first attaching a capture antibody to

the plate surface (indirect ELISA). Adapted by *life Technologies's* guide: Overview of ELISA.

The most commonly used detection enzymes are horseradish peroxidase (HRP) and alkaline phosphatase (AP). A large selection of substrates is available for both and the choice depends on the required assay sensitivity and the instrumentation available for detection. Though not as sensitive as fluorescent or chemiluminescent substrates, chromogenic ELISA substrates, like the one selected for this work, allow direct visualization and enable kinetic studies to be performed.

### **2.3.4 Microdot array**

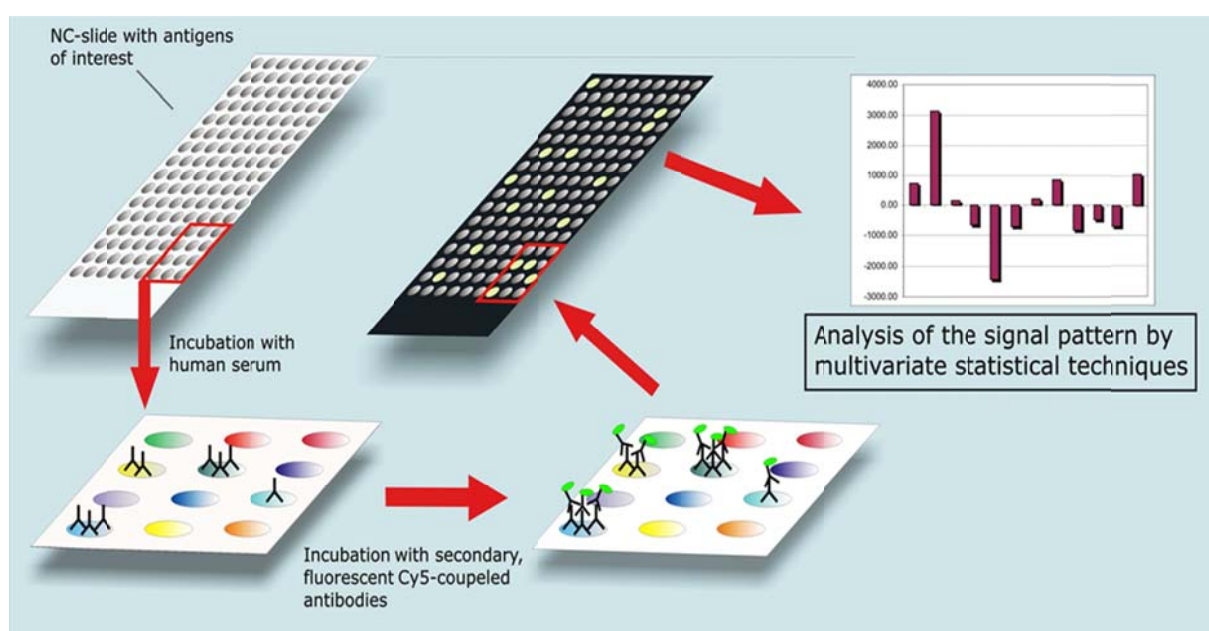
When working with protein-carbohydrate interactions, appropriate assays must be established to detect and amplify the outcome, since they are generally based on weak, noncovalent binding events. Inspired by the above described ELISA, a series of high-throughput methods have been formulated to track and quantify the interactions and activities of proteins and glycans. Among these techniques, the protein and glycan microarrays proved to be particularly useful to analyze, on large scale, the binding of specific antibodies to carbohydrates or proteins immobilized directly on the surface (glycan array) or via *de novo*-designed peptide scaffolds (protein array).

The first microarray was the DNA based;<sup>155</sup> from it, protein microarray has evolved starting from antibody microarrays<sup>156</sup> and nowadays they are frequently used in diagnostics, proteomics, antibody characterization and treatment development, together with glycan microarrays.<sup>157</sup>

The main advantage of a microarray is the large number of samples and conditions that can be tested simultaneously. Each one of the analytes is immobilized separately on the solid support, and each selected condition (for example different concentration) can be repeated even hundreds of times within the same experiment. The solid support on which the assay is performed consists of a small slide (generally glass, but also beads or a nitrocellulose membrane) on which a high number of analyte microdrops are spotted. The surface should possess maximal binding properties while maintaining the structural properties of the analyte, to guarantee the success of binding. To be able to host the analyte while preventing denaturation and displaying minimal non-specific binding, the solid support must be pre-treated with an immobilizing agent, such as amines, aldehydes or epoxies. The droplets are

then spotted on the slide with a pre-defined pattern by a printing robot. When the printing is complete, the ligands or interacting molecules are added on selected spots with the help of a microwell, which prevents cross-contamination between samples. The ligands can be fluorescently labeled for direct analysis or another fluorescent binding molecule can be added to the first for indirect examination. The analysis of the large amount of resulting data (means of fluorescence proportional to the binding of the ligand to the analyte) is generally facilitated by software for statistical calculations.

As for the ELISA, there are several variations of the basic procedure, and also in this case the indirect detection has been preferred for the work presented in this thesis. An example of the procedure is offered by Figure 18.

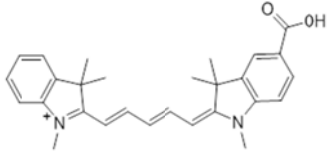
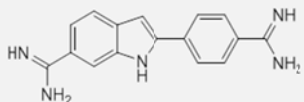
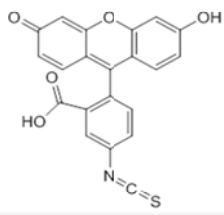
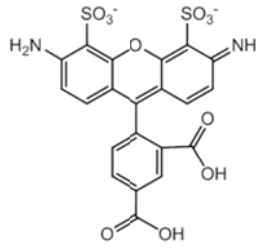
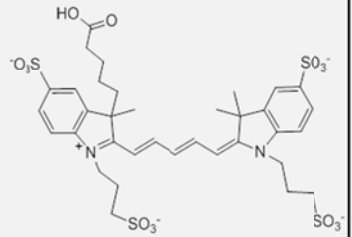


**Figure 18.** Workflow and principle of protein microarray technology. Adapted from “*Protein Microarray techniques*”, Copyright © 2008 PD Dr. Grus.

### 2.3.5 Employed fluorophores

The use of fluorophores in the work presented in this thesis was crucial in all biological assays. Fluorophores have been employed to follow the destiny of the peptides after binding to the cell, to detect the occurrence of the interactions between peptides and receptors, and to quantify these interactions. Numerous fluorophores will be encountered while reading this work. Table 2 summarizes their properties and uses.



Name	Structure	Excitation/Emission wavelength (nm)	Employment
<b>DM-IDCC COOH</b> here called <b>Cy5</b>		650/670	Coupled on peptide N-termini for flow cytometry and fluorescence microscopy.
<b>DAPI</b>		350/470	Used in fluorescence microscopy to stain cell nuclei (binds strongly to A-T rich regions of DNA).
<b>FITC</b>		490/525	Used in fluorescence microscopy for passive cell staining (E. coli) and coupled to ConA receptor.
<b>DyLight 488</b>	Structure undisclosed Patent # US 20140206099 CAS 1620475-58-2	493/518	Used in microdot assay for antibody binding quantification (coupled to secondary antibody).
<b>AlexaFluor 555</b>		555/565	Used in fluorescence microscopy for cell membrane staining (coupled to secondary antibody).
<b>AlexaFluor 647</b>		650/668	Used in microdot assay for antibody binding quantification (coupled to secondary antibody).

**Table 2.** Fluorophores used in the context of this thesis. The table reports name, structure, excitation/emission wavelength and employment.



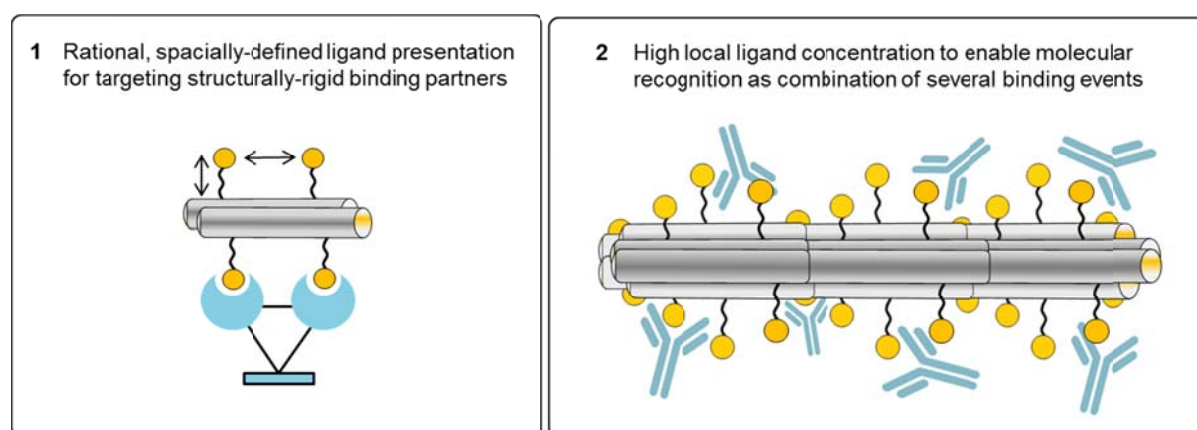
# 3

**Synthetic Glycopeptides as Multivalent Scaffolds for Carbohydrates:**  
from receptor targeting to vaccines exploiting sugar-protein interactions

## MOTIVATION AND OBJECTIVES

Peptides show great pharmaceutical potential for therapeutics and diagnostics, and represent a suitable scaffold due to their functionalities and folding properties. They are made even more attractive by the possibility of engineering them to display biologically-relevant ligands for interaction with specific binding partners. These characteristics are crucial for the development of efficient drug delivery systems, offering the necessary multifunctionality, architecture and solubility. The aim of this study is the development of coiled-coil peptides as drug-delivery systems, tools for diagnostics, and carriers for antigens, for which they are suitable based on the specificity of target recognition and their versatility. One of the main challenges for a peptide chemist is to design molecules characterized by sufficient specificity towards their targets, suitable stability and appropriate formulation. In this study, this challenge has been faced by exploiting the peptide conformation to bias binding towards particular macromolecules. The rational design of the coiled-coil scaffolds employed in this work has been addressed towards the presentation of carbohydrate ligands for selective drug delivery and multivalent antigen presentation.

This thesis is focused on two main concepts that exploit the use of the coiled coil as a scaffold: 1) the rational presentation of ligands on specific positions along a dimeric coiled-coil scaffold and with a specific orientation towards their binding partners; 2) the multivalent presentation of ligands on a highly oligomeric, fiber-forming coiled-coil scaffold to increase the chance of rebinding. A representation of these two concepts is given in Figure 19.



**Figure 19.** The two main concepts explored in this thesis. Grey cylinders: coiled coil-based scaffold; yellow circles: ligands. 1: tailored presentation of ligands to a rigid, dimeric binding partner (blue spheres). 2: multivalent presentation of ligands to specific antibodies (blue “Y”-shaped structures).

# 4

**Synthetic Glycopeptides as Multivalent Scaffolds for Carbohydrates:**  
from receptor targeting to vaccines exploiting sugar-protein interactions

## **PROJECT 1:**

# **TAILORED PRESENTATION OF CARBOHYDRATE LIGANDS ON A COILED-COIL SCAFFOLD FOR ASYALOGLYCOPROTEIN RECEPTOR TARGETING**

- 4.1 An overview of ASGPR
- 4.2 Glycopeptide library design and nomenclature
- 4.3 Glycopeptide synthetic strategy
- 4.4 Glycopeptide library structural characterization
- 4.5 Analysis of hepatocyte-mediated uptake
- 4.6 Inhibition studies of ASGPR-mediated endocytosis
- 4.7 Analysis of the internalization of selected glycopeptides within HepG2 cells
- 4.8 Summary and outlook I
- 4.9 Experimental procedures I

Targeted drug delivery is a method to deliver a beneficial molecule in a localized manner, minimizing side effects and reducing fluctuation in circulating drug levels. The main disadvantages connected to this technique are high costs and synthetic difficulties deriving from the production of the scaffold and the conjugation of the cargo. This project offers a step towards the resolution to these problems by evaluating the reliability of a dimeric coiled-coil peptide to be used as precision tool for the rational presentation of carbohydrate ligands to the asialoglycoprotein receptor (ASGPR).

A dimeric coiled-coil peptide has been exploited as a molecular ruler to display galactose ligands at different, well-defined distances between each other and via spacers of different length. An accessible synthetic strategy has also been proposed. The target macromolecule, the liver-expressed asialoglycoprotein receptor (ASGPR), binds the terminal galactose of glycoproteins and internalizes these via endocytosis for blood clearance purposes. Since binding can only occur when the ligands are presented at the correct distance and with a specific orientation, successful targeting of this receptor would represent a new approach to a more accessible targeted drug-delivery system.

#### 4.1 AN OVERVIEW OF ASGPR

ASGPR represents a rare example of an organ-specific receptor with high selectivity towards its binding partner.<sup>158</sup> These characteristics make it an ideal target for the selective delivery of biologically-active molecules to the liver, where it is mainly expressed.<sup>159, 160</sup> It is estimated that each human hepatocyte contains 100,000-500,000 such binding sites.

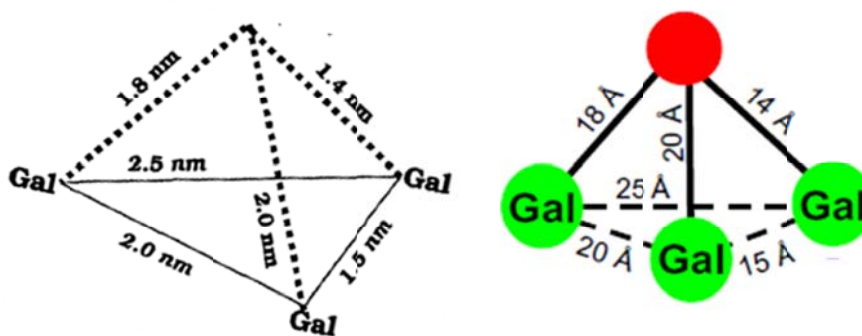
The ASGPR is a carbohydrate-binding protein which has the main physiological role of blood clearance from D-galactose-terminating glycoproteins when sialic acid is removed from complex *N*-linked oligosaccharides. However, the affinity of the receptor for D-*N*-acetylgalactosamine-terminating proteins and neoglycoproteins derivatized with this monosaccharide is substantially higher than for those terminating in galactose.<sup>161</sup>

Beyond the preferential binding for these carbohydrate residues, binding to the ASGPR depends strongly on the valency of the oligosaccharide; that is, mono-, bi-, tri-, and tetra-antennary galactose-terminal oligosaccharides bind with higher affinities, with dissociation constants of  $10^{-3}$ ,  $10^{-6}$ ,  $5 \times 10^{-9}$  and  $10^{-9}$  M, respectively, suggesting a close arrangement of at least three galactose binding sites.<sup>162</sup> Moreover, compounds within the same valency group exhibit a considerable range of affinities, suggesting that the inter-galactose distances are also important determinants. The importance of the inter-ligand distance and their relative spatial

orientation suggests that the receptor subunits must be held together in a rigid configuration to produce a complementary arrangement of binding sites.

The ASGPR is therefore thought to be triantennary, with one carbohydrate-recognition domain per subunit.<sup>163</sup> Composed of two homologous subunits, H1 and H2, only the structure of the H1 domain has been resolved.<sup>164</sup> Due to their specific binding, highly selective for the type and orientation of the carbohydrate, the lack of the complete receptor crystal structure represents a major obstacle in the use of the ASGPR in the delivery of drugs/genes to the liver.

Using Gal and GlcNAc-conjugated polymers, Lee *et al.* established an ASGPR ligand-based model, which is considered to be the most reliable one to date.<sup>165</sup> Khorev *et al.* also give a particularly clear adaptation of it.<sup>166</sup> This model and its adaptation can be found in Figure 20.



**Figure 20.** Optimal spatial arrangement of galactose residues in a triantennary glycan for the interaction with the ASGPR. On the left: Lee's original model. On the right: Khorev's adaptation (red sphere represent the construct; green spheres are galactose moieties (Gal)).

According to Lee's model, the ideal candidate for targeting the ASGPR displays three binding partners arranged in a pyramid branching from a common point and delimiting a scalene triangular base. The model was constructed based on a number of di- and tri-valent glycopeptides with terminal galactose residues.<sup>167, 168</sup> The structure with the highest ASGPR affinity was the one allowing a maximal inter-sugar separation of  $\approx 25 \text{ \AA}$ . Greater distances do not improve binding.

ASGPR activity requires  $\text{Ca}^{2+}$  ions (optimal concentration: 0.1-2 mM),<sup>169</sup> and it is therefore classified as a C-type lectin. The residues responsible for the interaction with the  $\text{Ca}^{2+}$  ions are the same ones that coordinate binding to the hydroxyl groups of the saccharide.<sup>170</sup> Following binding, the glycoprotein is internalized via endocytosis and directed to the lysosome for degradation. So-called "recycling receptors", like ASGPR, usually carry their ligands only to

the endosomes, where an acid-induced conformational change causes dissociation of the complex; the released ligands are degraded, while the receptors are sorted into recycling vesicles that bring them back to the plasma membrane for reuse.<sup>171</sup> These receptors have the characteristics of efficient high-turnover uptake systems.

## 4.2 GLYCOPEPTIDE LIBRARY DESIGN AND NOMENCLATURE

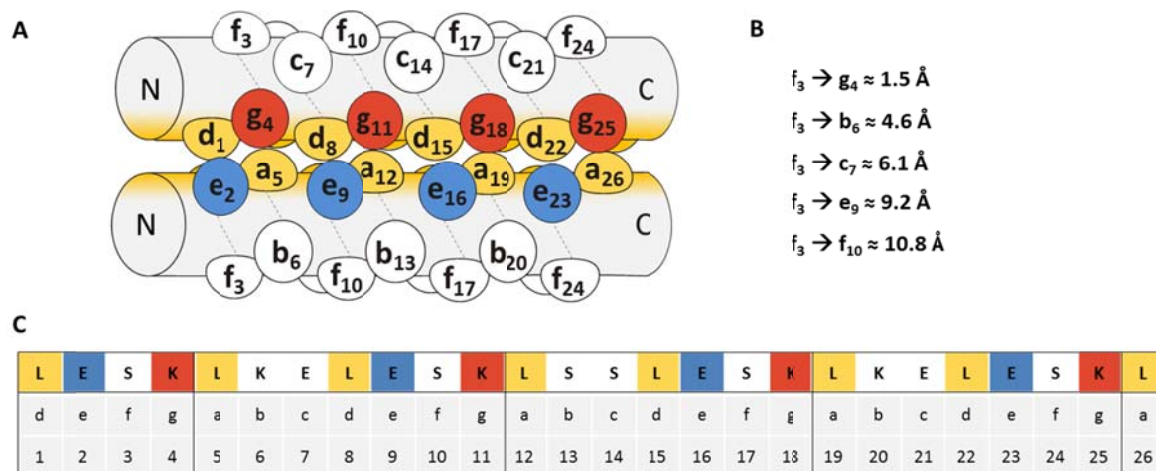
To optimize the fit between the presented galactose moieties and the ASGPR target, the coiled-coil scaffold must be able to carry one or more ligands with specific orientation and optimal spatial distribution. The coiled-coil peptide was chosen as scaffold since it offers the possibility of positioning the carbohydrates at specific distances along the helical axis. The folding principles of the coiled coil allowed us to create a glycopeptide library to verify efficiency and specificity of this motif in delivering carbohydrate ligands to a structurally rigid receptor. The components of the glycopeptide library vary in the number of displayed galactose moieties, their position on the peptide sequence and the space between the peptide backbone and the carbohydrate.

The parent coiled coil selected as scaffold was the 26-amino acid peptide of sequence H<sub>2</sub>N–LESKLKELESKLKELESKLKELESKL–OH, here named CCP. This peptide was previously designed in the group of Prof. Kokschi and was described in a report by Falenski *et al.*<sup>102</sup> CCP allows for the exclusive formation of coiled-coil dimers under physiological conditions and it can be extensively modified in solvent-exposed positions without altering its oligomerization state. Its sequence is composed of three and a half heptad repeats and contains four amino acids in the *f* position and a total of six amino acids in the *b* and *c* positions of the helical wheel. Modifications at these locations produce negligible structural variations, since these amino acids do not take an active part in folding. Thus, ten potential positions are available for the conjugation of ligands.

The structural simplicity and regularity of CCP (and of coiled coils in general) allow for the placement of each amino acid and conjugated ligand at predefined distances. For instance, two complete helical turns, required to allocate the seven residues of the coiled-coil heptad repeat, have a length of 10.8 Å (based on modeling studies). From this value, it is simple to calculate the distances between any two amino acid residues on the helix, considering that each one advances the helix by 1.5 Å along the main axis.<sup>172</sup> Such structural consistency translates into the possibility of a tailored delivery of ligands at pre-defined positions. The



distances between the  $\alpha$ -carbons, together with the sequence and a simplified cartoon of CCP, can be found in Figure 21.



**Figure 21.** The coiled-coil peptide CCP. A: The position of each residue composing the heptad repeats of CCP within the helix. B: Distances between the most solvent-exposed amino acids of the first heptad repeat ( $f_3$ ) and the other solvent-exposed residues of the heptad. C: Sequence, denomination within the heptads and number relative to the position in the sequence.

To be able to cover the longest possible distance within CCP main axis, we decided to construct a glycopeptide library bearing up to three galactose moieties (see sequence details below). To offer variety in the inter-galactose separation as well, we chose to place the carbohydrates either relatively close to each other, with a minimum of  $\approx 11$  Å, or as far apart as possible, at  $\approx 33$  Å. The glycopeptides bearing one galactose moiety only present it via conjugation of the sugar on position 3 of the sequence. When two galactose moieties are displayed, they are placed either on positions 3 and 10 or on position 3 and 24. Glycopeptides with three galactose ligands carry them in a short distance combination, on positions 3, 10 and 17, or in a long distance combination, on positions 3, 13 and 24. The distances between the sugar moieties on the scaffold CCP (based on modeling studies) are summarized in Table 3.

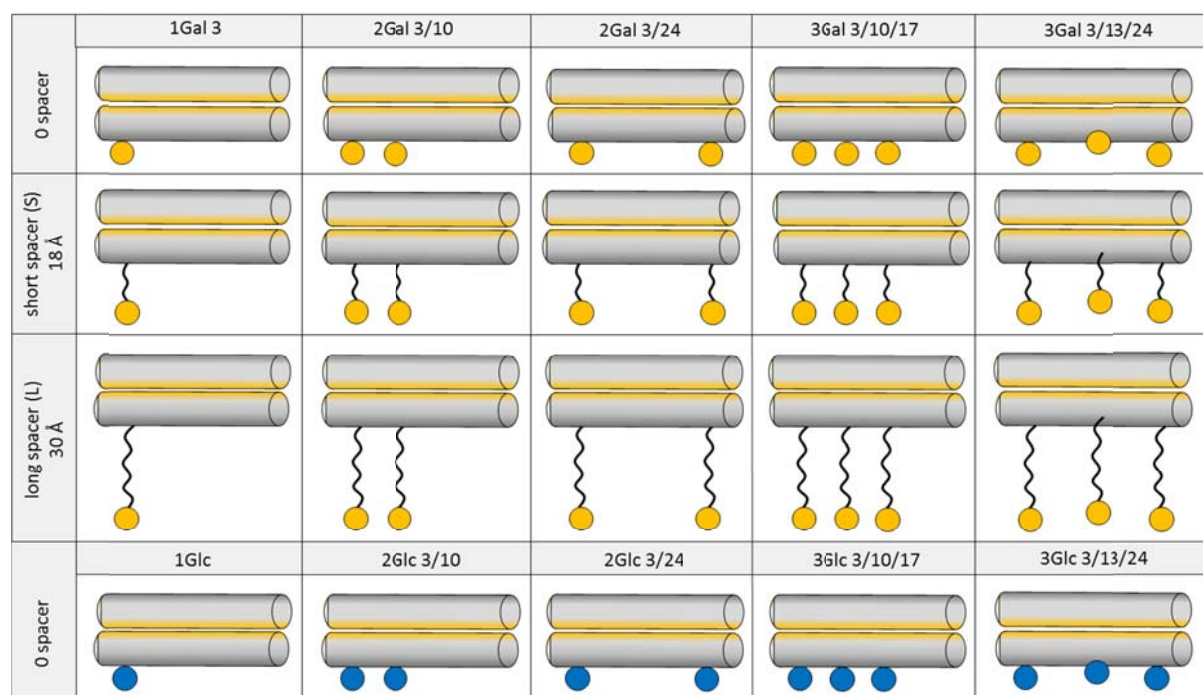
Position	S <sub>3</sub> -S <sub>10</sub>	S <sub>3</sub> -S <sub>13</sub>	S <sub>10</sub> -S <sub>17</sub>	S <sub>3</sub> -S <sub>24</sub>	S <sub>13</sub> -S <sub>24</sub>
Distance	$\sim 13$ Å	$\sim 19$ Å	$\sim 11$ Å	$\sim 33$ Å	$\sim 16$ Å

**Table 3.** Positions on the CCP sequence chosen to bear the galactose moieties and relative distances.

All sugar-bearing positions, except position 13, represent position *f* of the helical wheel, since it is well known to be the most-solvent exposed and the least involved in folding. The modification on position 13 occurs at the *b* position of the helical wheel, also solvent exposed, chosen to enable placement of three carbohydrate ligands equidistantly. This means that the ligand at this position is presented with a shift of  $\approx 1.5 \text{ \AA}$  compared to the others. If we acknowledge the flexibility given by the amino acid side chain and the spacer, this difference can be considered negligible.

Not only the distance between carbohydrate binding sites must be taken into account, but also the depth of the binding pockets of the receptor. For this reason, it is necessary to vary also the distance between peptide backbone and presented ligand. We chose to include spacers of three different lengths, by conjugating the galactose moieties directly on the amino acid side chain (0), via an 18  $\text{\AA}$  spacer (S) or via a 30  $\text{\AA}$  spacer (L). To guarantee a sufficient number of negative controls, we included into the library five peptides decorated with up to three glucose moieties directly on the backbone. Glucose is a non-relevant sugar for our receptor of interest, and any uptake of these glycopeptides by the hepatocytes would not be ASGPR-mediated.

Figure 22 depicts the full 20-member glycopeptide library derived from the combination of number of sugars displayed, distance among the galactose moieties and length of the spacer through which they are coupled on the peptide backbone.



**Figure 22.** Peptide (gray cylinders) library variants with up to three galactose (yellow circles) or glucose (blue circles) moieties. Columns: position of the sugars on the peptide; rows:

length of the spacer connecting peptide backbone and carbohydrates. The representation is a simplification of the structure: in fact, the ligands are presented from both helices.

The glycopeptides of the library are named according to the type and number of sugar moieties displayed (Gal or Glc), length of the spacer (0, S or L) and position along the helical axis (3, 3/10, 3/24, 3/10/17 or 3/13/24). Moreover, for synthetic purposes, the CCP sequence has been in some cases modified, replacing serine with lysine on the ligand-bearing positions with short or long spacer. The reasons and justifications for this choice will be examined in section 4.3. Nomenclature, sequence and description of the glycopeptide library are reported in Table 4. Noteworthy is the presence of a fluorophore (Cy5) at the N-terminal of each peptide. The dye has been coupled to be able to analyze the interaction of the library components with hepatocytes.

Name	Sequence	Description
1Gal-0 3	Cy5-HN-LES <sub>3</sub> KLKELESKLSSLESKLKELESKL-OH	1 Gal on position S <sub>3</sub> , no spacer
2Gal-0 3/10	Cy5-HN-LES <sub>3</sub> KLKELES <sub>10</sub> KLSSLESKLKELESKL-OH	2 Gal on positions S <sub>3</sub> and S <sub>10</sub> , no spacer
2Gal-0 3/24	Cy5-HN-LES <sub>3</sub> KLKELESKLSSLESKLKELES <sub>24</sub> KL-OH	2 Gal on positions S <sub>3</sub> and S <sub>24</sub> , no spacer
3Gal-0 3/10/17	Cy5-HN-LES <sub>3</sub> KLKELES <sub>10</sub> KLSSLES <sub>17</sub> KLKELESKL-OH	3 Gal on positions S <sub>3</sub> , S <sub>10</sub> , and S <sub>17</sub> , no spacer
3Gal-0 3/13/24	Cy5-HN-LES <sub>3</sub> KLKELESKL <sub>13</sub> SLESKLKELES <sub>24</sub> KL-OH	3 Gal on positions S <sub>3</sub> , S <sub>13</sub> , and S <sub>24</sub> , no spacer
1Gal-S 3	Cy5-HN-LE <b>K</b> <sub>3</sub> KLKELESKLSSLESKLKELESKL-OH	1 Gal on position K <sub>3</sub> , short spacer
2Gal-S 3/10	Cy5-HN-LE <b>K</b> <sub>3</sub> KLKELE <b>K</b> <sub>10</sub> KLSSLESKLKELESKL-OH	2 Gal on positions K <sub>3</sub> and K <sub>10</sub> , short spacer
2Gal-S 3/24	Cy5-HN-LE <b>K</b> <sub>3</sub> KLKELESKLSSLESKLKELE <b>K</b> <sub>24</sub> KL-OH	2 Gal on positions K <sub>3</sub> and K <sub>24</sub> , short spacer
3Gal-S 3/10/17	Cy5-HN-LE <b>K</b> <sub>3</sub> KLKELE <b>K</b> <sub>10</sub> KLSSLE <b>K</b> <sub>17</sub> KLKELESKL-OH	3 Gal on positions K <sub>3</sub> , K <sub>10</sub> , and K <sub>17</sub> , short spacer
3Gal-S 3/13/24	Cy5-HN-LE <b>K</b> <sub>3</sub> KLKELESKL <b>K</b> <sub>13</sub> SLESKLKELE <b>K</b> <sub>24</sub> KL-OH	3 Gal on positions K <sub>3</sub> , K <sub>13</sub> , and K <sub>24</sub> , short spacer
1Gal-L 3	Cy5-HN-LE <b>K</b> <sub>3</sub> KLKELESKLSSLESKLKELESKL-OH	1 Gal on position K <sub>3</sub> , long spacer
2Gal-L 3/10	Cy5-HN-LE <b>K</b> <sub>3</sub> KLKELE <b>K</b> <sub>10</sub> KLSSLESKLKELESKL-OH	2 Gal on positions K <sub>3</sub> and K <sub>10</sub> , long spacer
2Gal-L 3/24	Cy5-HN-LE <b>K</b> <sub>3</sub> KLKELESKLSSLESKLKELE <b>K</b> <sub>24</sub> KL-OH	2 Gal on positions K <sub>3</sub> and K <sub>24</sub> , long spacer
3Gal-L 3/10/17	Cy5-HN-LE <b>K</b> <sub>3</sub> KLKELE <b>K</b> <sub>10</sub> KLSSLE <b>K</b> <sub>17</sub> KLKELESKL-OH	3 Gal on positions K <sub>3</sub> , K <sub>10</sub> , and K <sub>17</sub> , long spacer
3Gal-L 3/13/24	Cy5-HN-LE <b>K</b> <sub>3</sub> KLKELESKL <b>K</b> <sub>13</sub> SLESKLKELE <b>K</b> <sub>24</sub> KL-OH	3 Gal on positions K <sub>3</sub> , K <sub>13</sub> , and K <sub>24</sub> , long spacer
1Glc-0 3	Cy5-HN-LES <sub>3</sub> KLKELESKLSSLESKLKELESKL-OH	1 Glc on position S <sub>3</sub> , no spacer
2Glc-0 3/10	Cy5-HN-LES <sub>3</sub> KLKELES <sub>10</sub> KLSSLESKLKELESKL-OH	2 Glc on positions S <sub>3</sub> and S <sub>10</sub> , no spacer
2Glc-0 3/24	Cy5-HN-LES <sub>3</sub> KLKELESKLSSLESKLKELES <sub>24</sub> KL-OH	2 Glc on positions S <sub>3</sub> and S <sub>24</sub> , no spacer
3Glc-0 3/10/17	Cy5-HN-LES <sub>3</sub> KLKELES <sub>10</sub> KLSSLES <sub>17</sub> KLKELESKL-OH	3 Glc on positions S <sub>3</sub> , S <sub>10</sub> , and S <sub>17</sub> , no spacer
3Glc-0 3/13/24	Cy5-HN-LES <sub>3</sub> KLKELESKL <sub>13</sub> SLESKLKELES <sub>24</sub> KL-OH	3 Glc on positions S <sub>3</sub> , S <sub>13</sub> , and S <sub>24</sub> , no spacer

**Table 4.** Peptide library nomenclature including sequence and description. The amino acid residues in bold indicate the positions bearing the ligands.

### 4.3 GLYCOPEPTIDE SYNTHETIC STRATEGY

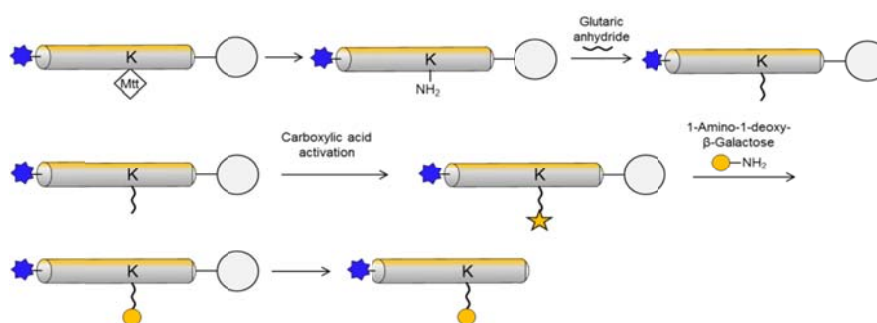
To reduce synthetic costs and complexity in the production of the glycopeptide library, a new accessible synthetic strategy was developed in the context of this thesis: the synthesis of the glycopeptide library was carried out completely on solid phase. We named this strategy “all-on-solid-phase (AOSP)” synthesis. Apart from this common feature, different synthetic approaches were employed according to the distance between backbone and monosaccharide, as schematized in the figures below (modifications are shown only on one of the two helices of the dimer).

In the case of the peptides without spacer (Figure 23), Fmoc-Ser-(*O*-beta-D-galactose-pentaacetate)-OH (Fmoc-Ser-Gal(Ac)) or Fmoc-Ser-(*O*-beta-D-glucose-pentaacetate)-OH (Fmoc-Ser-Glc(Ac)) was synthesized prior to SPPS and introduced at selected positions during SPPS. The chosen dye, Cy5, was coupled on the N-terminus as the last synthetic step, and the acetyl protecting groups of the sugars were removed after peptide cleavage from the resin:



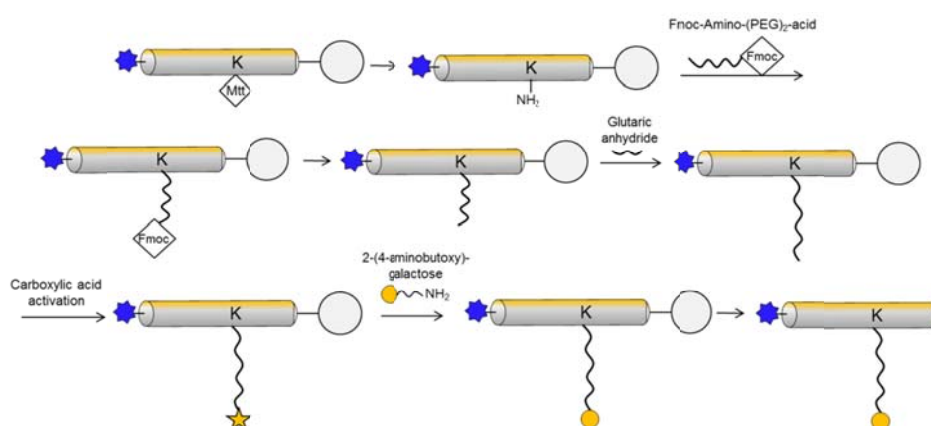
**Figure 23:** Synthesis of glycopeptides with no spacer. Gray circles: resin for SPPS; gray cylinders: peptides; yellow circles: galactose/galactose-derivatives; blue star: Cy5.

When the short linker was required (Figure 24), we replaced serine with N-methyltrityl (Mtt)-protected lysine on those positions selected to carry the carbohydrate ligand. The Mtt protecting group is removed under mild acidic conditions that ensure that the peptide remains anchored to the resin and all other amino acids remain protected. The side chain of these lysine residues was used to couple 1-Amino-1-deoxy- $\beta$ -Galactose by means of glutaric anhydride. Cy5 was coupled as above:



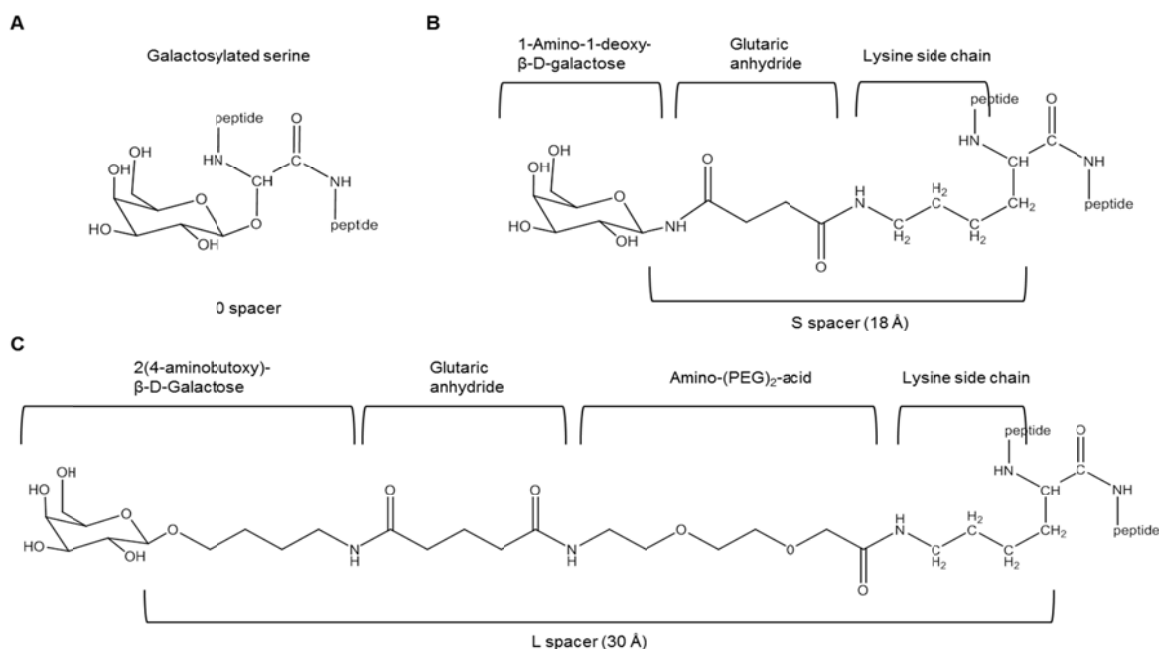
**Figure 24:** Synthesis of glycopeptides with short spacer. Grey circles: resin for SPPS; gray cylinders: peptides; yellow circles: galactose/galactose-derivatives; blue star: Cy5; yellow star: activation site.

In the case of peptides with the longer spacer (Figure 25), the first step was the completion of peptide synthesis with Mtt-protected lysine, for the further coupling of the sugar moieties, and with Cy5 at the N-terminus. In this instance, Fmoc-amino-(PEG)<sub>2</sub>-acid was introduced on selected lysine residues to gain a longer spacer. After Fmoc removal, the functionality was switched from –NH<sub>2</sub> to –COOH via glutaric anhydride. The acid was then activated on resin and aminobutoxy-galactose was coupled on solid phase:



**Figure 25:** Synthesis of glycopeptides with long spacer. Gray circles: resin for SPPS; gray cylinders: peptides; yellow circles: galactose/galactose-derivatives; blue star: Cy5; yellow star: activation site.

Below (Figure 26) the complete structure of the spacers is reported.

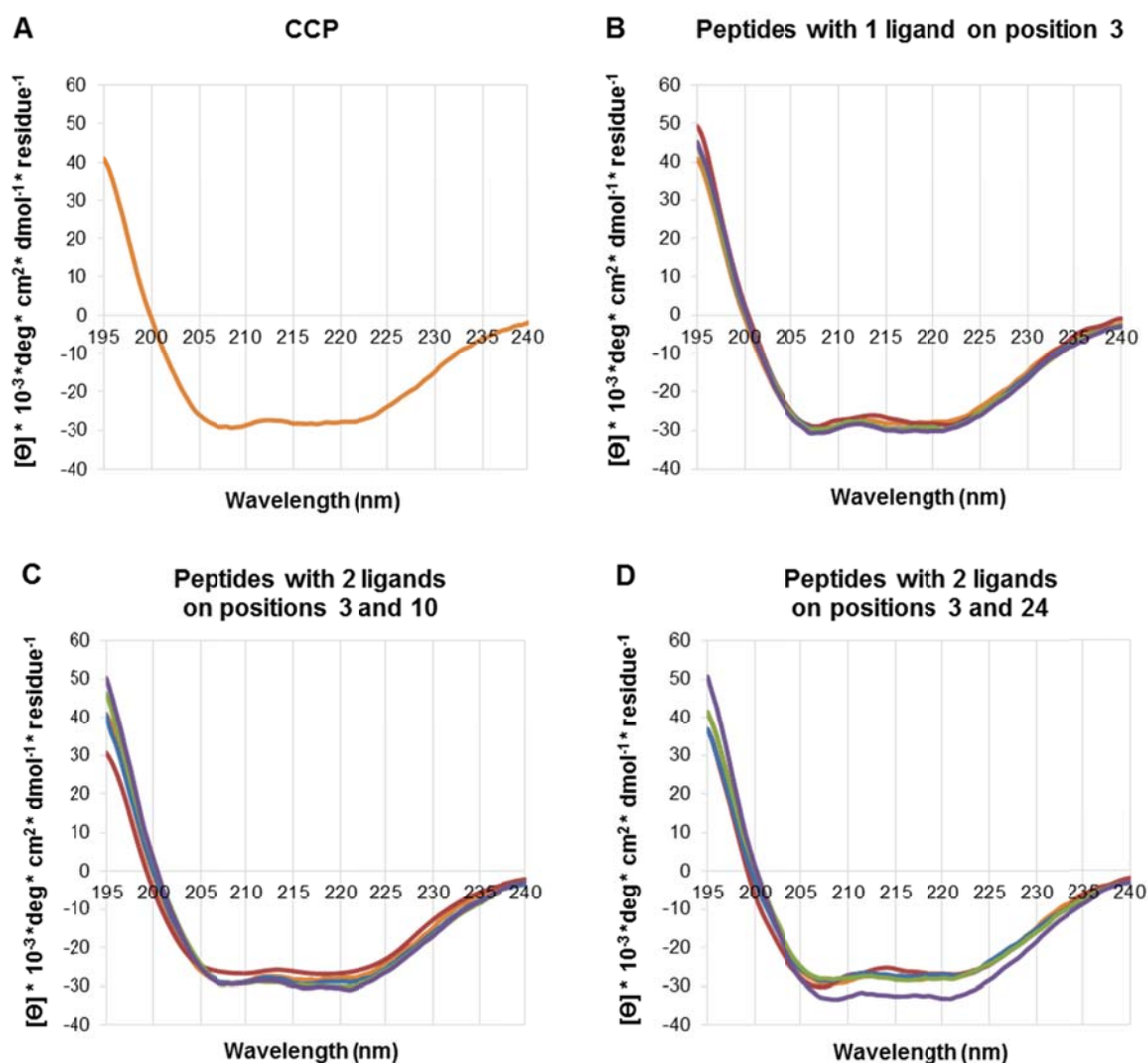


**Figure 26.** Chemical composition and structure of the spacers. A: Galactosylated serine; B: 18 Å spacer; C: 30 Å spacer.

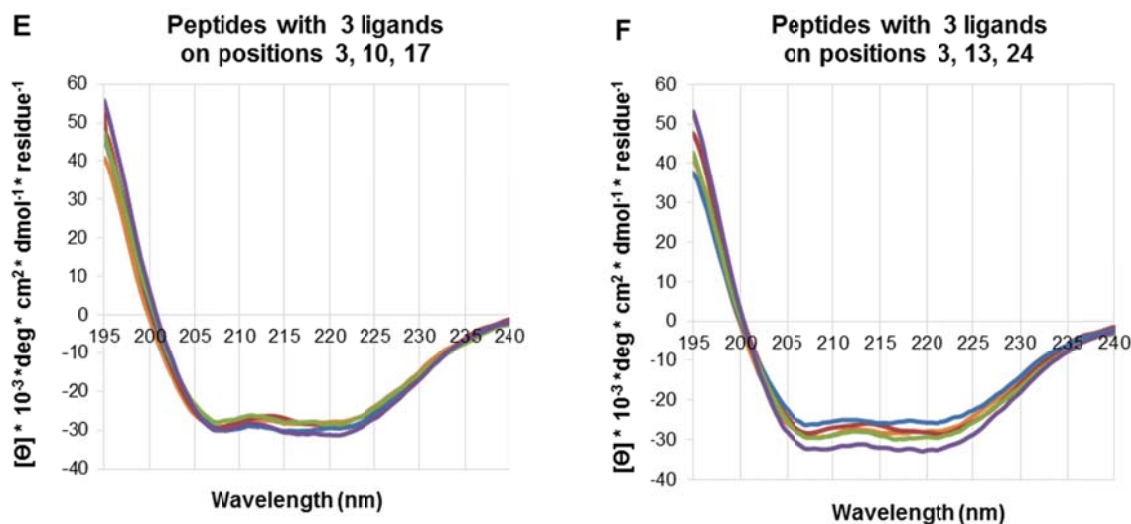
#### 4.4 GLYCOPEPTIDE LIBRARY STRUCTURAL CHARACTERIZATION

In the study by Falenski *et al.*, CCP was shown to be structurally stable and consistent, even when loaded with up to six galactose moieties on the peptide side chains. In the current study, the structural consistency among all components of the library must be investigated, since new variables have now been included (spacers, type and position of the sugars, primary sequence). Our interpretation of the peptide scaffold's reliability in delivering ligands is strictly connected to the structural consistency: only if all the members of the library prove to be  $\alpha$ -helices and show comparable degrees of ellipticity can we assume the above distribution of the ligands on the scaffold.

The glycopeptide secondary structures were therefore analyzed by means of CD spectroscopy and compared with the parental CCP to determine changes in conformation. Figure 27 reports the CD spectra of the glycopeptide library divided according to number and position of the sugar moieties.







**Figure 27.** Glycopeptide library CD spectra. A: Parental peptide CCP; B: Controls and peptides with one galactose, coupled at position 3; C: Controls and peptides with two galactoses, coupled to position 3 and 10; D: controls and peptides with two galactose moieties, coupled to position 3 and 24; E: Controls and peptides with three galactose moieties, coupled to positions 3, 10 and 17; F: Controls and peptides with three galactose moieties, coupled to positions 3, 13 and 24. CD spectra of different peptides are represented according to the following color code. Orange: CCP; red: Glc-functionalized peptides; blue: Gal-functionalized peptides, no spacer; green: Gal-functionalized peptides, short spacer; purple: Gal-functionalized peptides, long spacer.

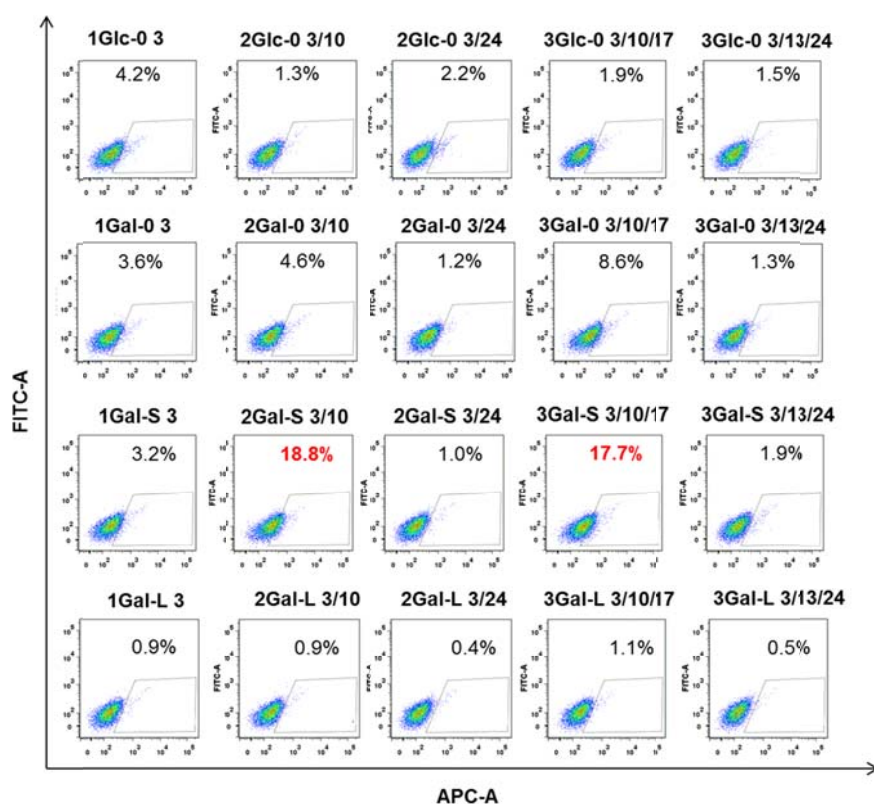
The CD spectrum of CCP is shown in panel A of Figure 27. Two minima at 208 nm and 222 nm, respectively, and a maximum at 195 nm are clearly distinguishable. The wavelengths corresponding to the signals and the relative intensities are attributable to the typical CD spectrum of an  $\alpha$ -helical peptide (see section 2.2.1).

Panels B to F compare the secondary structure of the library components with the CCP parent. All peptides, including the Glc-functionalized controls, show the typical CD profile of an  $\alpha$ -helical secondary structure, with a degree of ellipticity comparable to CCP. This means that, under the tested conditions, the introduced modifications have negligible impact on the ability of CCP to maintain the folding dictated by its sequence. Furthermore, the addition of ligands close to the N or C-terminus, where one might expect the highest helix fraying,<sup>173</sup> does not reveal a destabilizing influence on the coiled-coil structure, when compared to the parental CCP. Thus, we can assume that it is possible to achieve a tailored-presentation of ligands via this coiled-coil scaffold, and that the distances between the displayed sugar moieties can be predicted as described above.

#### 4.5 ANALYSIS OF HEPATOCYTE-MEDIATED UPTAKE

The interaction of the glycopeptide library with the ASGPR was screened by means of flow cytometry using HepG2 hepatic cells. HepG2 cells derive from a human hepatocellular carcinoma cell line. Due to their robust morphological and functional differentiation, they are a suitable model for studying intracellular trafficking. Glycopeptides bearing the fluorescent dye Cy5 were incubated with the hepatocytes for certain amounts of time at different concentrations. After flow cytometry, the outcome of the investigation is a percentage of cells within the population that, having interacted with the glycopeptides, now bear the fluorophore and emit a certain fluorescence signal. This allows us to identify the peptides that preferentially interact with the HepG2 by comparing the relative percentage of fluorescent cells and by using as specific controls the Glc-conjugated peptides.

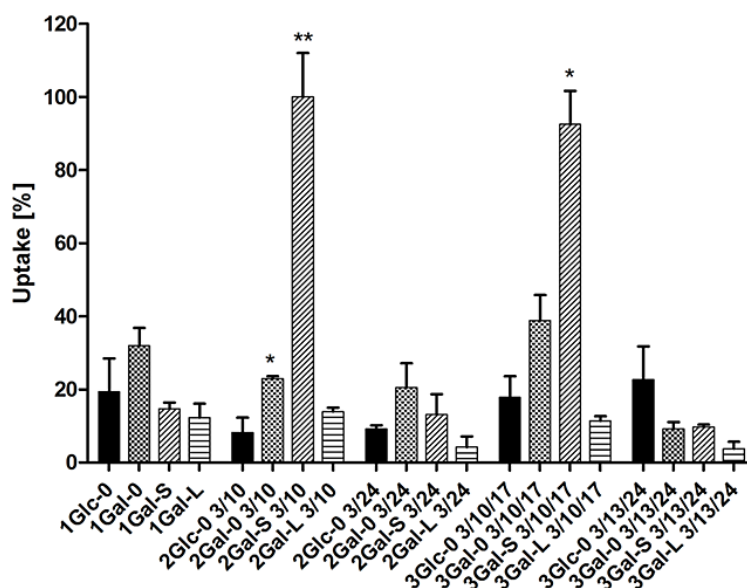
The density plots in Figure 28 report the results of the library screening.



**Figure 28.** ASGPR-mediated uptake of the glycopeptide library at a peptide concentration of  $0.01 \mu\text{M}$ . Each plot is relative to the uptake of one single peptide and depicts the percentage of Cy5-fluorescent cells on the y-axis and the light scattering on the x-axis. The colors indicate the number of cells (warm color: high cell density; cold color: low cell density).

Normalizing the results obtained from the library screening to the average uptake of the best binder, the histogram shown in Figure 29 was obtained:





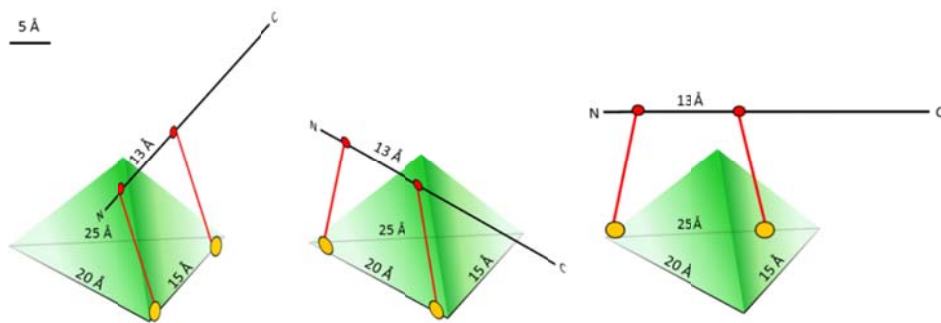
**Figure 29.** Normalized percentage of uptake of the glycopeptide library by HepG2 cells. Values were determined between each conjugate and the respective Glc-functionalized control peptide using the unpaired Student's *t* test (\* $p < 0.05$ , \*\* $p < 0.01$ ).

It appears evident that it is possible to identify two best binders for ASGPR-mediated uptake: 2Gal-S 3/10 and 3Gal-S 3/10/17. Both peptides present the galactose moieties at  $\approx 12$  Å from each other and contain the 18 Å (S) spacer. Slightly higher cell uptake compared to the respective glucose-conjugated controls was also detected for peptides 1Gal-0 3, 2Gal-0 3/10 and 3Gal-0 3/10/17. In all these cases, the sugars are coupled directly to the peptide backbone and both 2Gal-0 3/10 and 3Gal-0 3/10/17 glycopeptides present the galactose moieties at  $\approx 12$  Å from each other, as in the case of the best binders. By contrast, the peptides with galactose moieties coupled to the amino acids on positions 3/13 and 3/13/24, with or without spacer, exhibit uptake similar to the respective glucose-conjugate controls, and thus do not specifically bind to ASGPR on HepG2 cells. Furthermore, all galactose-functionalized peptides carrying the 30 Å (L) spacer exhibit the lowest cellular uptake. Negligible cell uptake was also observed for the monovalent galactose-conjugates.

Analyzing these results in light of the ligand-based ASGPR model published by Lee, a few additional observations can be made:

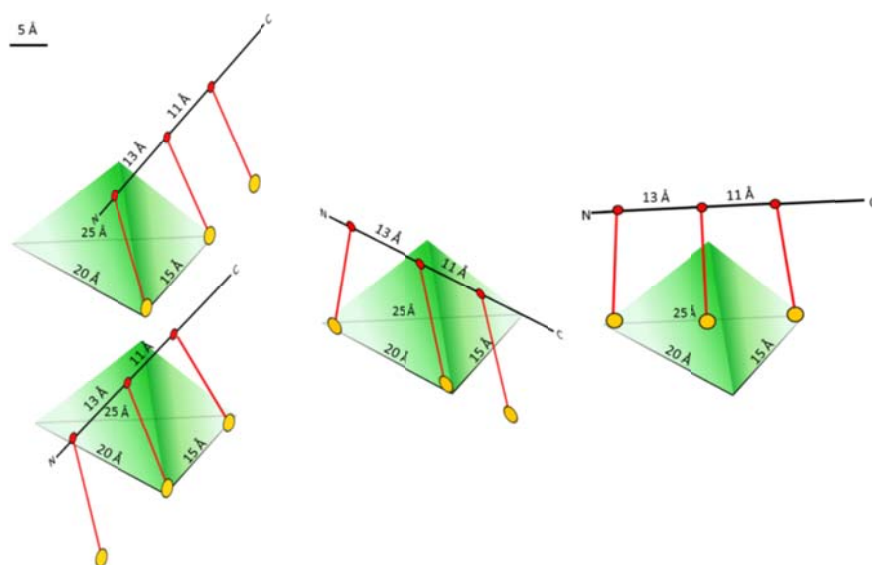
1. The two best binders are very similar, the only difference lying in the number of presented galactose moieties. Nevertheless, the addition of a third ligand does not improve uptake by the hepatocytes. An explanation to this behavior could be that, due to their intrinsic

geometry, both peptides target only two of the three carbohydrate binding pockets of the ASGPR. Details can be found in Figures 30 and 31.



**Figure 30.** Model of the interaction between 2Gal-S 3/10 with the binding pockets of the ASGPR. Pyramid: reproduction of Lee's model of the ASGPR binding pockets; black line: peptide backbone; red circles: positions on the peptide bearing the modifications; red line: S spacer (18 Å); yellow circle: galactose.

The glycopeptide 2Gal-S 3/10 can simultaneously interact with two pockets situated on the shorter side of the triangular base (first and second images within Figure 30). The two binding sites on the longest side of the triangle cannot be reached even considering the flexibility offered by the spacer (image on the right within Figure 30).



**Figure 31.** Model of the interaction between 3Gal-S 3/10/17 with the binding pockets of the ASGPR. Pyramid: reproduction of Lee's model of the ASGPR binding pockets; black line: peptide backbone; red circles: positions on the peptide bearing the modifications; red line: S spacer (18 Å); yellow circle: galactose.

Despite presenting three galactose moieties, only two of the ligands are optimally presented for interaction with the ASGPR. The shortest side of the triangular base can potentially be reached by both galactose at positions 3/10 and at positions 10/17 (two images on the left within figure 31). Due to the flexibility of the spacer, the binding pockets along the second shortest side of the triangular bases may also be reached by the sugars at positions 3/10, but likely not by those on positions 10/17 (middle image in Figure 31). Moreover, it is important to note that the galactose moieties at positions 3 and 17 are likely able to reach the binding pockets located along the longest side of the triangular base.

Despite being able to only interact with two of the here depicted binding pockets, the peptides 2Gal-S 3/10 and 2Gal-S 3/10/17 offer three possible combinations that could be used for recognition by ASGPR.

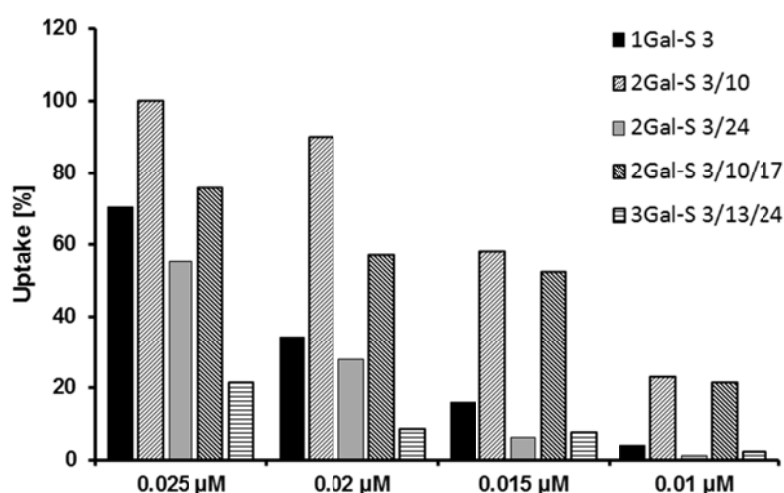
2. The glycopeptide bearing two ligands at positions 3/24,  $\approx 31.5 \text{ \AA}$  from each other, presents them too far apart to favor interaction with ASGPR, irrespective of the presence or length of the spacer. A better fit can be visualized only with the longest side of the triangular base according to which the carbohydrate-binding domains are distributed. Furthermore, ligands presented too far apart can promote cross-interaction with more than one receptor simultaneously, and this might facilitate dissociation, because the rebinding phenomenon that stimulate ligand uptake would be less likely to occur.

3. The glycopeptide 3Gal-S 3/13/24 displays galactose moieties at  $\approx 19 \text{ \AA}$  and  $\approx 16 \text{ \AA}$  from each other, respectively. The sugars on positions 3 and 13 are probably able to interact with the second shortest side of the triangle, while galactose moieties on positions 13 and 24 could better fit the binding pocket on the shortest side. If this were true, the same preferential binding by ASGPR as for peptides 2Gal-0 3/10 and 3Gal-0 3/10/17 would be expected. Since this is not the case, it must be assumed that the role played by the receptor surface density and clustering and the consequent likelihood of cross-linkage offered by the carbohydrates on positions 3 and 24 have a significant impact on the dissociation constant, as for the peptide 2Gal-S 3/24.

4. No peptides presenting the ligand with the L spacer ( $30 \text{ \AA}$ ) show preferential uptake when incubated with HepG2. This is likely due to the negative entropic contribution arising from the long, flexible spacer, which may disfavor ligand binding.<sup>174</sup>

5. The negligible non-specific binding of the glucose-conjugated peptides, defined as binding of the ligand to non-receptor domains, is expected to take place whenever the receptor is not saturable within the range of ligand concentrations used.<sup>175</sup>

To strengthen our results, we repeated the flow cytometry experiments using higher peptide concentrations. For the tested concentrations, 2Gal-S 3/10 and 3Gal-S 3/10/17 were the glycopeptides for which the HepG2 cells showed preferential binding, as reported in Figure 32.



**Figure 32.** Normalized percentage of uptake by HepG2 cells of the glycopeptides bearing galactose moieties via the short spacer. The histogram reports the results relative to four peptide concentrations, after two hours of incubation with HepG2 cells.

As depicted in Figure 32, the employment of higher peptide concentrations increases the chance of nonspecific binding, and the difference in uptake between our best peptide candidates and the relative controls become less pronounced, although they are still significant. This occurs because nonspecific binding is generally proportional to the concentration of the ligand and increases until receptor saturation. In most cases, the bulk of nonspecific binding represents some sort of interaction of the ligand with membranes. The molecular details are unclear, but nonspecific binding mostly depends on the charge and hydrophobicity of the ligand.

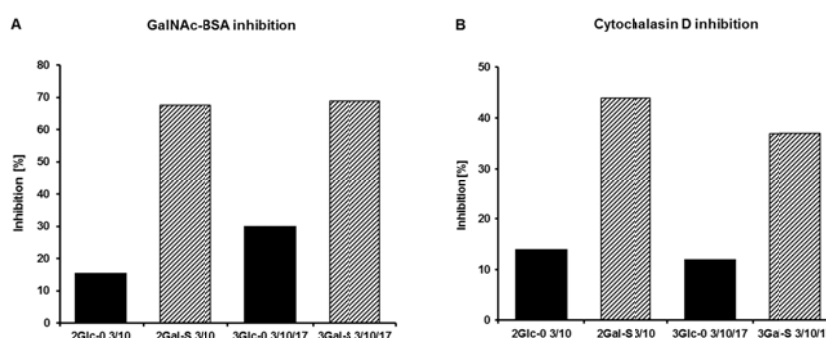
Noteworthy is the fact that, even at the higher concentration of 0.025 μM, the uptake of the peptides carrying the galactose moieties through the long linker is always relatively low. This suggests that the presence of the 30 Å-long PEG-based linker also limits the binding with other region of ASGPR and with the cell membrane. The reason for this behavior could lie in the high degree of freedom of movement offered by the spacer to the carbohydrates.

## 4.6 INHIBITION STUDIES OF ASGPR-MEDIATED ENDOCYTOSIS

As previously mentioned, the ASGPR is the most abundant receptor present on the surface of hepatocytes. This evidence, together with the results obtained via flow cytometry uptake studies by HepG2, suggests that the most successful glycopeptide binders of our coiled-coil library are taken up by the ASGPR. Nevertheless, to confirm that the uptake of the peptides 2Gal-S 3/10 and 3Gal-S 3/10/17 by HepG2 cells is indeed ASGPR-mediated, binding studies in the presence of a receptor-specific inhibitor must be conducted.

ASGPR-mediated endocytosis was inhibited by pre-treating the cells with GalNAc-BSA<sup>176</sup> or cytochalasin D.<sup>177</sup> GalNAc-BSA is produced by coupling GalNAc to bovine serum albumin (BSA) at multiple sites along its sequence. GalNAc, the preferred ligand for the ASGPR, is therefore multivalently presented to the receptor and, at high concentrations, saturates it. The binding of any other ligand would be then strongly reduced.

Cytochalasin D is an alkaloid produced by molds and is a potent inhibitor of actin polymerization. The actin cytoskeleton is required for receptor-mediated endocytosis in mammalian cells. Thus, inhibition of actin polymerization would block the action of the ASGPR and other hepatic membrane receptors. As a consequence, glycopeptides could not efficiently interact with it. Thus, we preincubated HepG2 cells with a 100-fold excess of GalNAc-BSA and cytochalasin D (10  $\mu$ M) before adding our glycopeptides (0.01  $\mu$ M) and carried out flow cytometry experiments with the two best binders and the respective controls. If the uptake of 2Gal-S 3/10 and 3Gal-S 3/10/17 is indeed ASGPR-mediated, we should observe binding inhibition higher than the one attributed to 2Glc-0 3/10 and 3Glc-0 3/10/17. The results of this study are summarized in Figure 33.



**Figure 33.** Inhibition of ASGPR-mediated endocytosis with GalNAc-BSA (A) and cytochalasin D (B) on selected glycopeptides. Patterned columns: 2Gal-S 3/10 and 3Gal-S 3/10/17; solid columns: 2Glc-0 3/10 and 3Glc-0 3/10/17. Values are presented as percentage of inhibition, defined as the reduction of conjugate uptake by cytochalasin D- or BSA-treated cells divided by uptake by untreated cells.

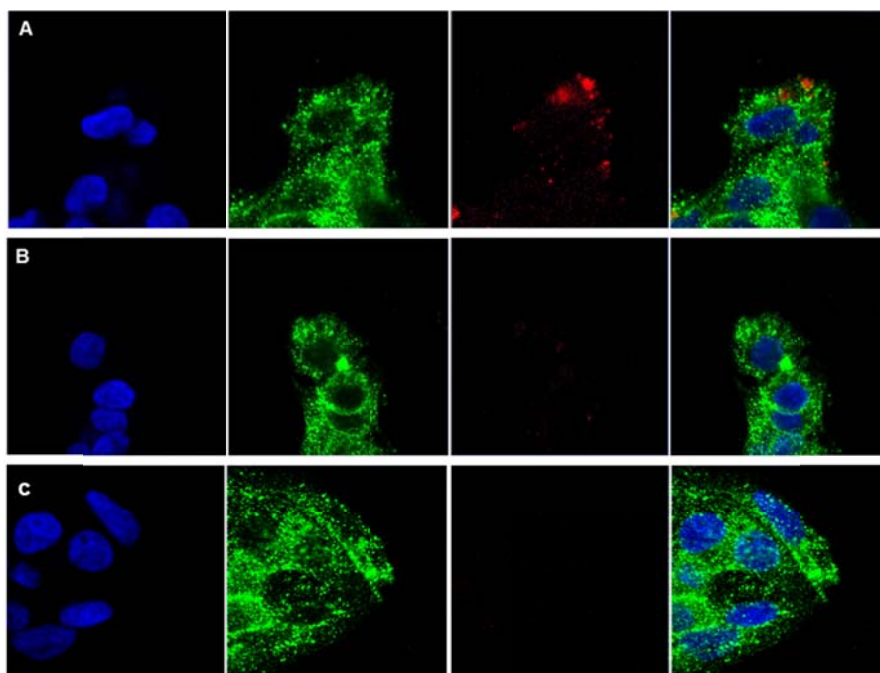
The uptake of the glycopeptides 2Gal-S 3/10 and 3Gal-S 3/10/17 undergoes a reduction of almost 70% when HepG2 cells are pre-treated with GalNAc-BSA. Inhibition relative to the control peptides 2Glc-0 3/10 and 3Glc-0 3/10/17 under the same conditions is instead marginal ( $\approx$  15% and 28%, respectively). These results strongly suggest that the uptake of the two glycopeptides 2Gal-S 3/10 and 3Gal-S 3/10/17 is ASGPR-mediated.

If it could be argued that the presence of the carrier molecule BSA in this assay might influence the recognition of the binding partners by the ASGPR, the confirmation of the result comes from the inhibition assays performed with cytochalasin D. Also in this case pre-incubation with the inhibitor leads to a greater decreased uptake of the Gal-functionalized peptides compared to their Glc-functionalized counterparts.

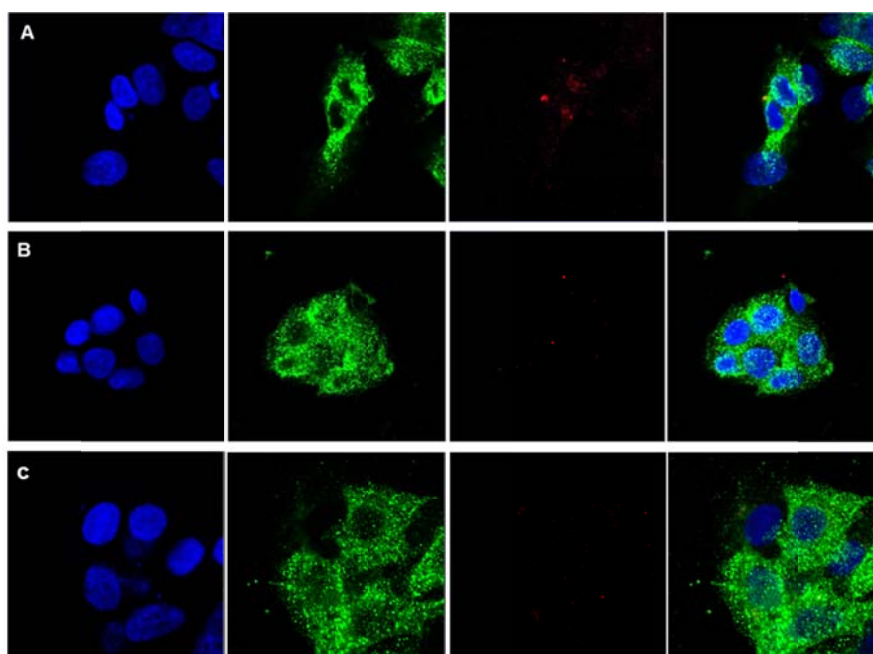
#### **4.7 ANALYSIS OF THE INTERNALIZATION OF SELECTED GLYCOPEPTIDES WITHIN HepG2 CELLS**

Flow cytometry enables quantitative determination of the interaction between the glycopeptide library and the hepatocytes. This measurement is relative to the cells that show fluorescence due to binding to the library members, but does not distinguish between peptides bound to the cell membrane and peptides that have actually been internalized within the cell. If the binding is ASGPR-mediated and the receptor does not undergo modifications that could interfere with its activity, the ASGPR should be able to internalize the bound glycopeptides by endocytosis and the fluorescently labelled peptides should be visible inside the cell. Therefore, to visualize and confirm the results obtained via flow cytometry experiments, HepG2 cells were incubated with the two best coiled-coil binders (2Gal-S 3/10 and 3Gal-S 3/10/17), with their respective controls (2Glc-0 3/10 and 3Glc-0 3/10/17) and with the glycopeptides showing lowest binding affinity (2Gal-L 3/10 and 3Gal-L 3/10/17). The cells were then observed under the fluorescence microscope (FL), to validate the presence of Cy5-labelled peptides inside cells. Nuclei and cell membrane were stained with dyes with emission properties different from Cy5 to facilitate recognition of subcellular components. In particular, red fluorescence associated with Cy5 represents peptide distribution, blue fluorescence (DAPI) associates with HepG2 nuclei and green fluorescence indicates emission from the Alexa Fluor 555-conjugated secondary antibody (Ab) against an anti-ATPase primary Ab, present in the cell membrane and cytoplasm. In the figures below, the three fluorescence signals were examined individually (first three panels) and merged (fourth panel). The FL images were acquired at different peptide concentrations and after different

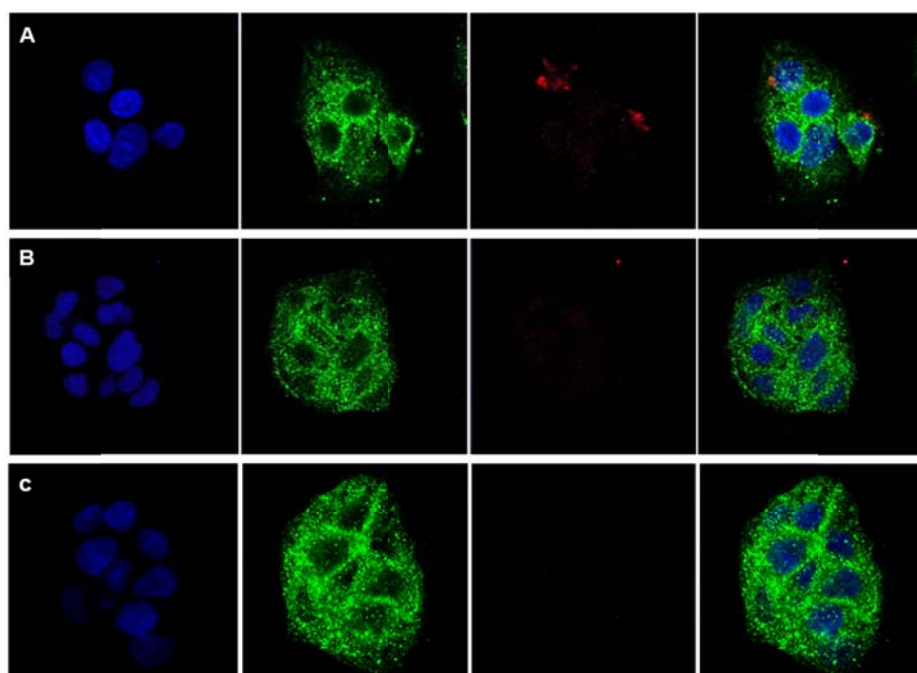
incubation times with HepG2 cells. Some representative images are shown in Figures 34-35, for the glycopeptides bearing two carbohydrate moieties, and Figures 36-37, for the glycopeptides with three carbohydrate moieties.



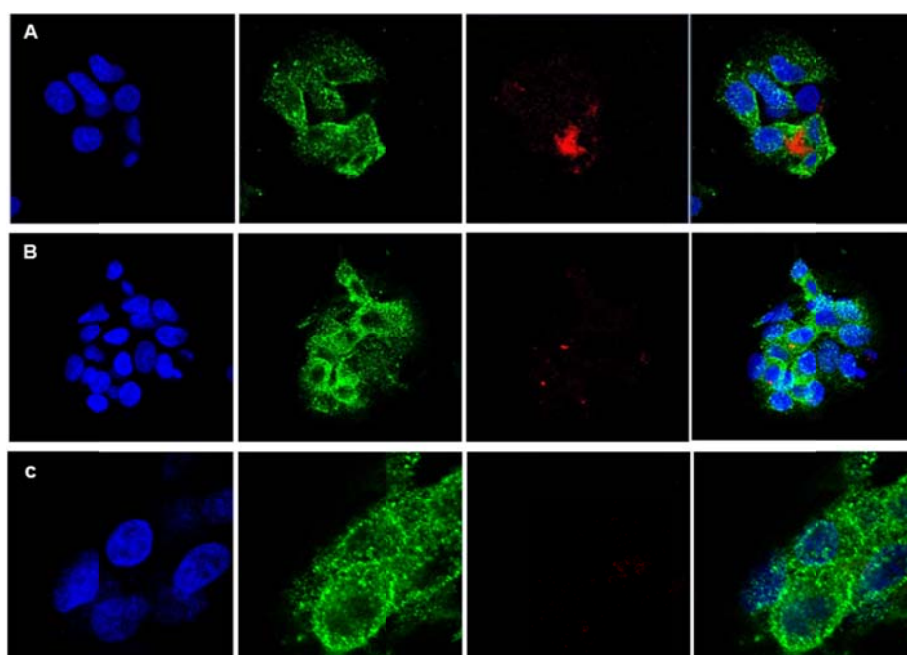
**Figure 34.** Fluorescence microscopy images of HepG2 cells after incubation with Cy5-labeled A: 2Gal-S 3/10; B: 2Glc-0 3/10; C: 2Gal-L 3/10. Images were taken after 2 hours incubation, at a peptide concentration of 0.1  $\mu\text{M}$ .



**Figure 35.** Fluorescence microscopy images of HepG2 cells after incubation with Cy5-labeled A: 2Gal-S 3/10; B: 2Glc-0 3/10; C: 2Gal-L 3/10. Images were taken after 30 minutes incubation, at a peptide concentration of 0.4  $\mu\text{M}$ .



**Figure 36.** Fluorescence microscopy images of HepG2 cells after incubation with Cy5-labeled A: 3Gal-S 3/10/17; B: 3Glc-0 3/10/17; C: 3Gal-L 3/10/17. Images were taken after 2 hours incubation, at a peptide concentration of 0.1  $\mu\text{M}$ .



**Figure 37.** Fluorescence microscopy images of HepG2 cells after incubation with Cy5-labeled A: 3Gal-S 3/10/17; B: 3Glc-0 3/10/17; C: 3Gal-L 3/10/17. Images were taken after 30 minutes incubation, at a peptide concentration of 0.4  $\mu\text{M}$ .



The Cy5-labeled glycopeptides 2Gal-S 3/10 and 3Gal-S 3/10/17 (panels A in figures 34-37) are clearly located inside the hepatocytes. At both peptide concentrations and incubation time of 30 minutes and 2 hours, the distribution of red fluorescence, which appears as small dots, suggests that the peptides are likely found within endocytosis vesicles. Moreover, the vesicles appear to be spread throughout the cytosol, sparing the nuclei, as expected for endocytic uptake.

Qualitatively, it appears that under the selected conditions endocytosis of the peptides 2Gal-S 3/10 and 3Gal-S 3/10/17 is more efficient than the respective Glc-controls 2Glc-0 3/10 and 3Glc-0 3/10/17 (panel B in Figures 34-37). Peptides presenting two galactose moieties attached to the long spacer, 2Gal-L 3/10 and 3Gal-L 3/10/17, show negligible internalization, at a level similar to the glucose controls, in accordance with the flow cytometry experiments (section 4.5).

The FM observations confirm and reinforce the results obtained by means of flow cytometry: among the distances and spacers selected to build this glycopeptide library, members bearing galactose moieties at an average of  $\approx 12 \text{ \AA}$  via a  $18 \text{ \AA}$  spacers are the most suitable coiled-coil carriers for targeting the ASGPR.

#### 4.8 SUMMARY AND OUTLOOK I

The work described here was published within the framework of this thesis as an original research article entitled “*Tailored Presentation of Ligands on Coiled Coil-Based Glycopeptides for Asialoglycoprotein receptor Targeting*”,<sup>178</sup> and it is a systematic study of the reliability of the coiled-coil motif in functioning as a scaffold for the well-defined presentation of ligands to biological macromolecules.

For this purpose, a 26-amino acid dimeric coiled-coil peptide was chosen as carrier molecule for the display of galactose moieties to the mammalian asialoglycoprotein receptor (ASGPR). The ASGPR is a carbohydrate-binding protein almost exclusively found on the cell membrane of hepatocytes and it exclusively binds galactose- and N-acetylgalactosamine-terminating proteins for blood clearance purpose; it internalizes them via endocytosis and delivers them to the lysosome before being recycled to the cell surface.

To target this receptor, we designed and synthesized a 20-member coiled-coil glycopeptide library in order to identify the ideal fit for the ASGPR. The members of the library differed in the number, position and distance of the galactose moieties from the peptide backbone. We conjugated one, two or three galactose moieties in three regimes: 1) direct coupling to the

peptide backbone; 2) coupling via an 18 Å linker; or 3) coupling via a 30 Å linker.

Furthermore, the carbohydrate moieties were placed at shorter or longer intervals along the scaffold backbone.

HepG2 hepatocytes were incubated with each library member, including glucose-conjugated negative controls. By means of flow cytometry, we identified the best binders that optimally present the galactose ligands, and they are in agreement with models from earlier literature reports regarding both sugar position and spacer length.

To confirm that the binding was indeed ASGPR-mediated, we inhibited the endocytic pathway by means of cytochalasin D, a microtubule polymerization inhibitor, and with GalNAc-BSA, a known specific ASGPR binder. These inhibition studies showed that the uptake of the best glycopeptide binders by HepG2 cells was inhibited to a much greater extent than the binding of the glucose-conjugated controls. This demonstrated that the uptake of these glycopeptides by hepatocytes is indeed ASGPR-mediated.

To determine whether uptake also led to internalization of the ASGPR binding partners or if instead binding to the cell membrane was unspecific, we observed the HepG2 cells under the fluorescence microscope after incubation with our glycopeptides. These fluorescence microscopy studies showed that the peptides are clearly present inside the cells, and therefore do become internalized via endocytosis.

Within this study it was demonstrate that the coiled-coil motif is an excellent scaffold for the tailored presentation of ligands, particularly when the ligands must be displayed in an optimal spatial arrangement in order to interact selectively and with high affinity with the target. This conclusion suggests that the applicability of the coiled-coil scaffold can be expanded to the targeting of numerous receptors for various applications:

### **1. Drug delivery**

To deliver drugs and biologically-active molecules selectively and precisely to a certain cell type is an urgent need in clinical therapeutics, since in many cases the avoidance of nonspecific interactions is crucial to preserve patient health. For example, this type of coiled-coil scaffold could be designed to present ligands for receptors that are exclusively present or overexpressed on cancer cells. A pro-apoptotic molecule could be attached to the scaffold at an appropriate site and, following uptake by a specific receptor, the carrier peptide and the drug would be internalized, leading to cell death.

### **2. siRNA delivery**

Gene expression can be modulated by RNA-interference techniques, in which a short interfering RNA (siRNA) is delivered into the cells. The siRNA sequence is complementary to the mRNA of the target gene and binding leads to its degradation. When gene therapy is required, it is essential to consider the problems posed by the delivery of nucleic acid into the cells, since their negatively-charged surface does not allow them to cross the cell membrane. A coiled-coil scaffold could be functionalized with a receptor-interacting ligand and with a siRNA of interest to deliver the cargo inside the cell by facile interaction with specific membrane receptors.

### 3. Medical imaging

Medical imaging refers to the group of techniques that aim to provide visual representation of the interior of the patient's body for clinical analysis and medical intervention. These methods generally refer to magnetic resonance, nuclear imaging, ultrasounds and optical imaging. The latter implies the use of a fluorophore. As the results reported within this thesis demonstrate, the coiled-coil motif optimally bears both fluorophore and a selected ligand for specific cell/tissue/organ interaction. The use of coiled-coil based peptides in medical imaging could have a great impact on *in vivo* diagnostics at the nanoscale, and for the efficient diagnosis and targeting of molecular markers of disease.

## 4.9 EXPERIMENTAL PROCEDURES I

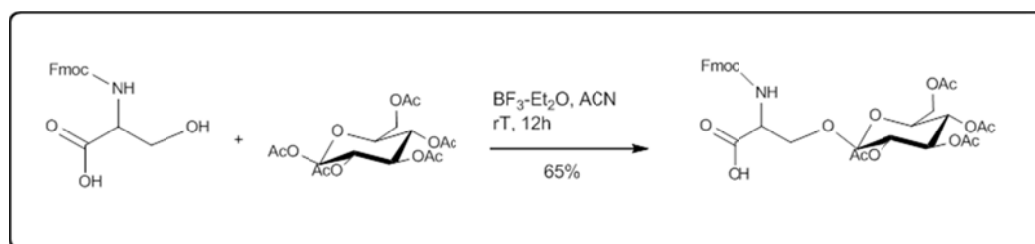
### Synthesis and purification of Fmoc-Ser-(O-beta-D-glucose-pentaacetate)-OH

The synthesis of Fmoc-Ser-(O-β-D-glucose-pentaacetate)-OH was carried out by dissolving β-D-glucose-pentaacetate (4 mmol) (Sigma, 285943) and vacuum-dried Fmoc-Ser-OH (1.2 eq.) (Iris Biotech, FAA1578) in dry acetonitrile (20 mL). The solution was placed on ice and BF<sub>3</sub>-Et<sub>2</sub>O (1 eq.) was added. The mixture was then left to slowly warm up to room temperature. An additional equivalent of BF<sub>3</sub>-Et<sub>2</sub>O was added after both 6 and 15 hours of stirring. After 20 hours, the reaction was shown by TLC to be complete and the mixture was diluted with DCM (30 mL) and sequentially washed with a 1 M HCl (3 x 30 mL) and H<sub>2</sub>O (2 x 30 mL). The organic phase was dried with MgSO<sub>4</sub> and the solvents were removed under reduced pressure. The crude product was purified by RP-HPLC using a gradient of 40%-100% MeOH in H<sub>2</sub>O to obtain 2.5 mmol (62.5 % yield) Fmoc-Ser-(O-beta-D-glucose-pentaacetate)-OH.

Fmoc-Ser-(O-beta-D-galactose-pentaacetate)-OH was purified with RediSep® Rf Reversed-phase, using a C18 column with average particle size: 40-63 microns, mesh: 230-400, average pore size: 60 Å, on a CombiFlash Rf, Teledyne Isco, Inc. Lincoln, NE, USA.

<sup>1</sup>H-NMR (250 MHz, CD<sub>3</sub>CN) δ 1.93 (s, 3H, Ac), 1.94 (s, 3H, Ac), 1.97 (s, 3H, Ac), 2.00 (s, 3H, Ac), 3.75-3.85 (m, 2H, Ser-βH, H-6), 4.05 (dd, 1H, J = 12.2 Hz, 2.4 Hz, Ser-βH), 4.11 (dd, 1H, J = 10.5 Hz, 4.3 Hz, H-6'), 4.25 (t, 1H, J = 6.4 Hz, Fmoc CHAr), 4.41 (m, 1H, Fmoc OCH<sub>2</sub>), 4.32-4.37 (m, 1H, Ser-αH), 4.61 (d, 1H, J = 8.0 Hz, H-1), 4.85 (dd, 1H, J = 9.7, 8.0 Hz, H-4), 5.01 (t, 1H, J = 9.7 Hz, H-2), 5.22 (t, 1H, J = 9.6 Hz, H-3), 5.80 (d, 1H, J = 7.9 Hz, NH), 7.35 (brt, 2H, J = 7.5 Hz, Ar), 7.43 (brt, 2H, J = 7.4 Hz, Ar), 7.67 (d, 2H, J = 7.2 Hz, Ar), 7.84 (d, 2H, J = 6.9 Hz, Ar).

MS (ESI-TOF, -mode). Compound with chemical formula C<sub>32</sub>H<sub>35</sub>NO<sub>14</sub>, calculated mass: 657,21; found: [M - H]<sup>-</sup> = 656.19.



**Scheme 1.** Reaction scheme and yield for the synthesis of Fmoc-Ser-(O-beta-D-glucose-pentaacetate)-OH

### Synthesis of 2(4-aminobutoxy)-β-D-Galactose

β-D-Galactose-pentaacetate (2.6 mmol) and 1.5 eq. 4-(Z-amino)-1-butanol (Aldrich, 95887) were dissolved in 15 mL of dry DCM. One equivalent of BF<sub>3</sub>-Et<sub>2</sub>O was added to this solution as it stirred on ice. The mixture was then left to slowly warm up to room temperature. An additional equivalent of BF<sub>3</sub>-Et<sub>2</sub>O was added after both 6 and 15 hours of stirring. After 20 hours, the reaction was shown by TLC to be complete. The mixture was treated with a saturated solution of NaHCO<sub>3</sub> (30 mL), washed with H<sub>2</sub>O (2 x 30 mL) and brine (3 x 30 mL) and dried with Na<sub>2</sub>SO<sub>4</sub>. The crude product (pale yellow) was concentrated and purified by silica gel column chromatography (2:1 Hex:EtOAc) to yield 1.8 mmol (69%) of the glycosylated intermediate.

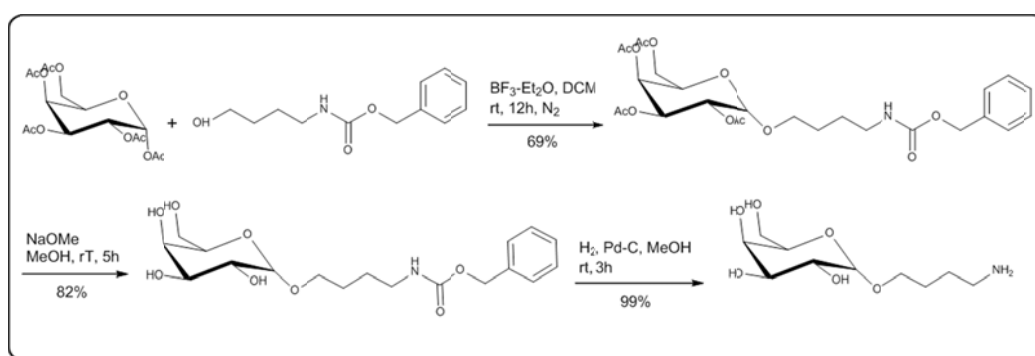
To remove the acetyl protecting groups, the glycosylated intermediate was dissolved in dry MeOH (5 mL) and 0.8 mmol 30% NaOMe (0.5 eq.) solution in MeOH was added dropwise. After stirring at room temperature for 4 h, the solution was treated with DOWEX 50WX8 50-100(H) resin and stirred until acidic. Upon filtration and solvent removal, 1.48 mmol (82% yield) of the deacetylated intermediate was obtained as a colorless oil.

In order to remove the carboxybenzyl protecting group, the deacetylated intermediate was dissolved in dry MeOH (15 mL) and catalytic Pd on activated carbon (0.1 eq.) was added. The suspension was purged with hydrogen gas for twenty minutes and then stirred under a hydrogen-atmosphere for 3 h. After filtration over celite, the final product was obtained as a colorless, viscous oil (99%).

<sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD) δ 1.58 (quint, 2H, J = 6.7 Hz, -CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>), 1.64 (quint, 2H, J = 6.7 Hz, -OCH<sub>2</sub>CH<sub>2</sub>-), 2.68 (t, 2H, J = 7.0 Hz, -CH<sub>2</sub>NH<sub>2</sub>), 3.29 (m, 1H, H-6), 3.44-3.51 (m, 3H, H-5, -OCH<sub>2</sub>-), 3.54-3.58 (m, 1H, H-4), 3.69-3.76 (m, 2H, H-2, H-6'), 3.88-3.93 (m, 1H, H-3), 4.20 (d, 1H, J = 7.3 Hz, H-1).

<sup>13</sup>C-NMR (500 MHz, CD<sub>3</sub>OD) δ 28.04 (-OCH<sub>2</sub>CH<sub>2</sub>-), 29.84 (-CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>), 42.06 (-CH<sub>2</sub>NH<sub>2</sub>), 62.47 (C-6), 70.28 (-OCH<sub>2</sub>-), 70.48 (C-4), 72.55 (C-2), 75.00 (C-3), 76.60 (C-5), 104.98 (C-1).

MS (ESI-TOF, +mode). Compound with chemical formula C<sub>10</sub>H<sub>21</sub>NO<sub>6</sub>; calculated mass: 251,14; found: [M + H]<sup>+</sup> = 252.15. Calculated mass + Na: 274.13; found: [M + Na]<sup>+</sup> = 274.13.



**Scheme S2.** Reaction scheme and yield for the synthesis of 2(4-aminobutoxy)-β-D-Galactose

### Glycopeptide library synthesis

Synthesis of the peptide library was performed with standard Fmoc/tBu chemistry, using Fmoc-protected amino acids purchased from Orpegen. Fmoc-Leu-NovaSyn TGA resin, 0.21

mmol/g amino acid loading, was purchased from Novabiochem. For the amino acids that did not carry any modification, the synthesis was carried out with 2-hour double couplings, using a SyroXP-I peptide synthesizer (Multi-SynTech GmbH), with a 5-fold excess of amino acid (relative to resin loading), TBTU/HOBt (5 eq.) and DIPEA (5 eq.). The addition of the Cy5-like dye at the N-terminus (DM-IDCC COOH, mivenion, 602200), containing a carboxylic acid for coupling, was achieved by repeating the reaction three times, using 1.5, 0.75 and 0.75 eq. of dye, respectively.

The addition of the galactose moieties and the spacers was performed using different strategies, according to the length of the spacer. When no spacer was required, Fmoc-Ser-Gal/Glc-pentaacetate was added by manual double coupling, with a three-fold excess of glycosylated amino acid, HATU (2.9 eq.) and DIPEA (6 eq.). For the peptides carrying the 18 Å and the 30 Å spacers, Mtt-protected lysine was incorporated at the positions of the sequence to which galactose would be linked (3, 10, 13, 17, 24). After coupling Cy5 as above, selective deprotection of the Mtt-protected lysine residues was performed.<sup>179</sup> The free lysine side chain was used for the coupling of the spacers. In the case of the 18 Å spacer, glutaric anhydride (five-fold excess) and a catalytic amount of DIPEA were used to switch the functionality from the amine of the lysine side chain to a carboxylic acid. After three hours, the acid was activated on resin with 1.2 eq. HATU and 2.4 eq. DIPEA for 5 minutes, and 1-amino-1-deoxy-β-D-galactose (2 eq.) (Sigma, A2267) was added. In the case of the 30 Å spacer, Fmoc-O<sub>2</sub>Oc-OH (amino-PEG-acid, Iris Biotech, 08575) was coupled to the deprotected lysine side chains using a four-fold excess of amino-PEG-acid, HATU (3.9 eq.) and DIPEA (8 eq.). The coupling was repeated twice and left to incubate for 6 hours. After removal of Fmoc from the amino-PEG-acid, glutaric anhydride was added and activated as previously described. 2(4-Aminobutoxy)-β-D-galactose was finally coupled on solid phase, repeating the reaction three times and using, respectively, 1, 0.75 and 0.75 eq., of building block, HATU and DIPEA.

Resin cleavage was performed using a solution of 95% TFA, 3% water and 2% TIS (5 mL) with shaking for 3 hours. The resin was washed with TFA and DCM and solvents were removed by evaporation. The peptides were precipitated in cold diethyl ether and collected by centrifugation.

Only for the peptides containing Fmoc-Ser-Gal/Glc-pentaacetate was a deacetylation reaction performed after cleavage. The peptides were dissolved in ACN and the minimum volume of MeOH required for complete dissolution. 10% vol. hydrazine (80% solution in water) was added and the mixture was stirred for 45 minutes. To quench, the flask was placed on ice and

a five-fold excess of AcOH was added and stirred for 10 minutes. The volatiles were removed under reduced pressure and the peptides purified via RP-HPLC.

### **Glycopeptide library purification**

Peptide purification was performed on a Knauer RP-HPLC. All runs were performed with a flow rate of 20.0 mL/min using a gradient of 15%-60% ACN (0.1% TFA) in Millipore H<sub>2</sub>O (0.1% TFA) over 30 minutes.

The collected fractions were analyzed by analytical RP-HPLC (LaChrom-ELITE (VWR)). All runs were performed with a flow rate of 1.0 mL/min using a gradient of 5%-70% ACN (0.1% TFA) in Millipore H<sub>2</sub>O (0.1% TFA) over 20 minutes. Data analysis was performed with EZ Chrom ELITE software.

### **Glycopeptide library concentration determination**

Peptide concentrations were determined using the extinction coefficient of the Cy5 dye in MeOH ( $150000 \text{ M}^{-1} \text{ cm}^{-1}$ , data obtained from mivenion). 50  $\mu\text{L}$  of the peptides previously dissolved in PBS buffer, pH 7.4, were added to 950  $\mu\text{L}$  MeOH and the absorbance at 650 nm was recorded by means of a Cary 50 Bio UV-spectrophotometer. The concentration of the stock solution (M) was calculated as follows:  $(\text{Abs}_{650 \text{ nm}} * \text{dilution factor})/150000$ .

### **CD spectroscopy**

All CD spectra were recorded on a JASCO-8-10 spectropolarimeter at 20 °C. The spectra were acquired in 0.1 mm path length quartz cuvettes. During all measurements a constant N<sub>2</sub> flush of 3.0 l/min was provided. Each obtained CD data set is the average of three measurements.

### **Flow cytometry-based cell-uptake studies**

HepG2 cells (ATCC® HB-8065), a human hepatocellular carcinoma-derived cell expressing ASGPR, were used for cell uptake studies with the carbohydrate-functionalized coiled-coil peptides. The cells ( $8 \times 10^4$ ) were seeded in 48-well plates using complete DMEM medium (supplemented with 10 % fetal calf serum, FSC, 100 U/mL penicillin, 100  $\mu\text{g}/\text{mL}$

streptomycin and 2 mM L-glutamine) and cultivated overnight. Cells were washed once with PBS and 100  $\mu$ L fresh FCS-free DMEM containing the carbohydrate-functionalized coiled-coil peptides (10 nM) was added. After two hours of incubation at 37°C, the cells were washed three times with ice-cold PBS supplemented with 0.5% BSA. Subsequently, the cellular uptake of the fluorescently labeled coiled-coil peptides was measured by flow cytometry using a FACSCanto II flow cytometer (BD Biosciences, San Jose, CA, USA). Data were analyzed with the FlowJo analysis software (Tree Star Inc., Ashland, OR, USA). For the inhibition of receptor-mediated endocytosis, cells were preincubated with cytochalasin D (10  $\mu$ M; Santa Cruz, Dallas, TX, USA) for one hour before addition of the carbohydrate-functionalized coiled-coil peptides.

### **Statistical analysis**

Statistical analyses were performed with the unpaired Student's t test. Data were analyzed using Prism software (GraphPad Software, La Jolla, CA, USA). A *p*-value of  $p < 0.05$  was considered statistically significant.

### **Confocal fluorescence microscopy**

HepG2 cells were left 24h in DMEM (PAN, P04-03500, supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 10% FCS and 100 U/mL penicillin/streptomycin), and seeded into 24-well plates ( $1 \times 10^5$  cells/well) with 12 mm-diameter round cover slides (Menzel GmbH).

Peptide stock solutions were prepared in sterile PBS and diluted in DMEM to reach concentrations of 0.1  $\mu$ M, 0.2  $\mu$ M and 0.4  $\mu$ M. Peptide solutions (100  $\mu$ L) were added to the HepG2 cells, left 10 minutes on ice to block unspecific binding and incubated for 30 minutes or 2 hours at 37 °C. The cells were washed one time with PBS and 2 times with BPBS, (PBS with 0.5% BSA) and fixation was carried out with 4% paraformalin, at room temperature.

From this point, the cells were kept protected from light throughout incubation. After washing (3x BPBS), 200  $\mu$ L of a solution of 100 mM glycine buffer in PBS (pH 7.4) was added for 10 minutes and the cells were washed again. For membrane staining, PMCA ATPase antibody (Thermo Scientific, 5F10) was diluted 1:400 in PBS, and the solution was added to the wells (200  $\mu$ L/well) for 1 hour at room temperature. After washing, 200  $\mu$ L of Alexa Fluor 555 goat anti-mouse IgG (life technologies, A21422) were added to the wells (dilution 1:200) for 1 hour at room temperature. Subsequently, the nuclei were stained with DAPI (Invitrogen,



21490, final concentration 500 nM, 200  $\mu$ L/well) for 5 minutes at room temperature, and the cells were washed again.

For observation under the microscope, each 12 mm-diameter round cover slide on which the cells were fixed was placed onto a microscope slide. Mounting medium (thiodiethanol buffer, 20  $\mu$ L) was added to each slide and the coverslips were sealed with nail polish. Fluorescent images were acquired with the instrument LSM 700 from Carl Zeiss, using three different lasers (405 nm, 555nm, 639 nm) for excitation and objective Plan Apochromat 63X/1.4 oil DIC. The images were processed with the software ZEN 2009 by Zeiss.



# 5

**Synthetic Glycopeptides as Multivalent Scaffolds for Carbohydrates:**  
from receptor targeting to vaccines exploiting sugar-protein interactions

## **PROJECT 2:**

# **A SELF-ASSEMBLING PEPTIDE SCAFFOLD FOR THE MULTIVALENT PRESENTATION OF ANTIGENS**

- 5.1 An overview of carbohydrate-based vaccines
- 5.2 Multivalent scaffold design
- 5.3 Glycoconjugate synthetic strategy
- 5.4 Evaluation of peptide library structural features
- 5.5 Multivalent presentation of ligands to selected binding partners
- 5.6 Summary and outlook II
- 5.7 Experimental procedures II

The stimulation of the immune response towards a specific target is the central concept of vaccination. One of the obstacles during vaccine production is determining the optimal concentration: this must be high enough to activate the immune system but low enough to avoid lateral effects. For this reason, various synthetic and natural scaffolds have been used to optimize the display of antigens in a multivalent fashion. In this way, the global concentration of the ligand remains low while the local concentration increases.

So far, self-assembling peptides have not been thoroughly investigated for this purpose. In this study, a *de novo*-designed fiber-forming peptide increases the local concentration of two antigens, simultaneously. The goal of this research is to evaluate whether this peptide could be a robust scaffold for the presentation of multiple ligands, mainly for diagnostic purposes and vaccination. In this work, the scaffold has been modified with a peptide epitope and a carbohydrate antigen and its ability to retain the parental supramolecular structure has been tested. The scaffold has been further evaluated for its capability to multivalently present both ligands to specific antibodies, thus enhancing recognition.

## **5.1 AN OVERVIEW OF CARBOHYDRATE-BASED VACCINES**

The activation of a protective immune response against any infectious agent depends initially on the identification, by the immune system, of an antigen as foreign. Immune-stimulatory compounds, or antigens, provoke one or more different types of reactions from the host in an effort to remove or disable the invading organism. The antigen may stimulate T-lymphocytes (T-cells), which provide cell-mediated immunity, or B-lymphocytes (B-cells), that initiate the synthesis and secretion of soluble antibodies into the bloodstream. The development of the body's protective immune response depends upon achieving a threshold level of stimulation of one or both of these systems.

Immunization by vaccination typically offers long-term protection since it activates both B-cells and T-cells, generating antibodies specific for a certain pathogen and allowing the development of immunological memory. Immune-response activation is obtained by employing an attenuated form of the pathogen or isolated polysaccharides or glycoproteins from the parasite cell wall or capsule. The resulting antibody production falls off after some days, but future exposure to the live, virulent organism will be promptly responded to with an even greater abundance of antibodies.<sup>180</sup> The main feature of carbohydrate-based vaccines is

the employment of polysaccharides or oligosaccharides as antigens.<sup>181</sup> The advantages of using glycans as antigens are numerous:

- Fine-tuned synthetic techniques, such as one-pot synthesis<sup>182</sup> and solid-phase synthesis,<sup>183</sup> allow for the synthesis of large quantities of pure sugars at much lower costs and higher yields than would be gained by the isolation of the same compounds from natural sources.
- Carbohydrates are highly antigenic, so they easily bind to lymphocytes and their receptors.<sup>184, 185</sup>
- In contrast to polypeptides, which are fully encoded by the corresponding gene, carbohydrate expression is not under direct genetic control but results from a dynamic interaction between the environment and a network of hundreds of genes;<sup>186</sup> thus, the evolution of resistance may be expected to occur less quickly than is the case for polypeptide epitopes.

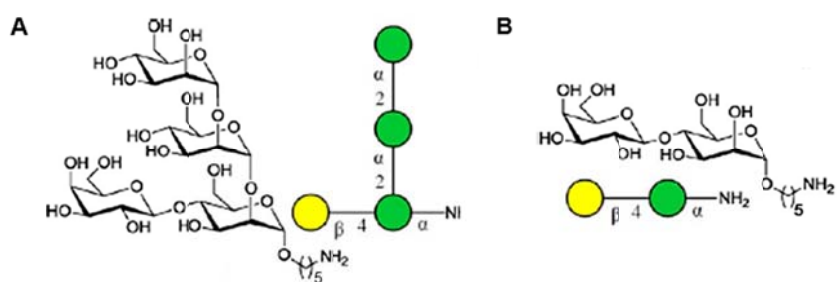
The main drawback in the employment of carbohydrates in vaccination is their poor immunogenicity, meaning that they are T-cell-independent antigens, and immunization solely with carbohydrates would not lead to immunological memory. Proteins are significantly more immunogenic than polysaccharides and, since carbohydrates are non-immunogenic haptens, they require conjugation with an epitope such as a protein before they can evoke a cell-mediated immunity. For this reason, when a carbohydrate-based vaccine is produced, T-cell epitopes are generally also employed.<sup>187</sup> In fact, although T-cells themselves do not secrete antibodies, they are frequently needed to assist in the stimulation of B-cells. In the production of a carbohydrate-based vaccine, the carbohydrate antigens are, therefore, generally conjugated to T-cell epitopes. The T-cell epitopes can be used in combination with an unrelated B-cell determinant to obtain significant quantities of antibody production against the B-cell antigen. The T-cell determinant is ingested and processed by the antigen presenting cells (APC), which will then present the epitope on their surface, in complex with the self molecules of the major histocompatibility complex (MHC). Helper T-cells can interact with this complex and activate specific B-cells, which will differentiate into plasma cells and secrete antibodies. Beside the immunological relevance of T-cell epitopes, their use as peptides, as opposed to larger proteins containing epitopes, may provide economic advantages, due to facile synthesis, as well as improved safety. To obtain higher efficacy and to increase the local concentration of the antigen, the antigen-epitope complex must be also conjugated to a carrier molecule. Therefore, there are in total four main elements to consider

in the synthesis of a carbohydrate-based vaccine: the selected carbohydrate antigen, the T-cell epitope, the carrier molecule and the conjugation method. These aspects are described in more detail in the following sections.

### 5.1.1 The carbohydrate antigen

The selection of the carbohydrate antigen is dictated by the particular infectious agent against which the immune response is to be stimulated. As mentioned in section 1.3.3, carbohydrate-based vaccines have been synthesized and/or commercialized for a large number of infectious pathogens. For example, vaccines against bacterial pathogens such as *Haemophilus influenzae* type b (Quimi-Hib),<sup>188</sup> *Streptococcus pneumoniae* (Prevnar),<sup>189</sup> and *Salmonella typhi* (VICPS)<sup>190</sup> are commercially available.<sup>191</sup> Also virus capsids are a subject of study in very recent vaccine development; among these, influenza virus A,<sup>192</sup> picornavirus<sup>193</sup> and human papillomavirus<sup>194</sup> vaccines are on the market. A number of studies also address the treatment and prevention of cancer, and mainly involve the MUC1<sup>195</sup> and Globo<sup>196</sup> antigens. Ultimately, a series of carbohydrate-based vaccines are currently under investigation for protozoan parasites, including malaria<sup>197</sup> and leishmaniasis.<sup>198</sup>

In the context of the study presented in this thesis, a diagnostically-relevant lipophosphoglycan (LPG) found on pathogenic *Leishmania* parasites was selected as the carbohydrate antigen. LPGs are major leishmanial cell surface glycoconjugates and play an important role in disease physiopathology. In particular, the tetrasaccharide cap Gal $\beta$ (1-4)- $\beta$ -[Man $\rho$ -(1-2)- $\alpha$ -Man $\rho$ -(1-2)- $\alpha$ ]-Man $\rho$ <sup>199</sup> plays a key role in the infection process. Recent studies conducted by our collaboration partners at the Max Planck Institute (MPI) of Colloids and Interfaces, led by Prof. Seeberger, have demonstrated that the most immunologically relevant carbohydrate antigen among *Leishmania* species is in fact the disaccharide  $\beta$ -Gal-(1 $\rightarrow$ 4)- $\alpha$ -Man, which constitutes the backbone repeat unit of most of the LPGs.<sup>200, 201</sup> Due to its desirable antigenic properties and synthetic accessibility, especially compared to the complex tetrasaccharide cap, this disaccharide, here referred to as Man-Gal, was chosen to be the carbohydrate-antigen subject of this study. Figure 38 shows the structure and saccharide symbols relative to both of these glycans.



**Figure 38.** Leishmania parasite antigens structure and color-code representation. A: tetrasaccharide cap Galp(1-4)-β-[Manp-(1-2)-α-Manp-(1-2)-α]-Manp. B: antigenic disaccharide β-Gal-(1→4)-α-Man, here named Man-Gal. The amino-pentanyl linker connected to the OH of the anomeric carbon of the mannose is the one introduced by our collaborators from the MPI.

### 5.1.2 The T-cell epitope

The epitope is the part of the antigen recognized by the immune system. T-cell epitopes, in particular, are generally short portions of antigenic proteins, composed of 13-17 amino acid residues. T-cells mainly recognize processed fragments of protein antigens (linear peptide epitopes) on antigen-presenting cells. Theoretically, any of these protein fragments could be selected as T-cell epitopes, but only about 2% have the correct amino acid sequence that allows them to strongly bind to MHC. The immunogenicity of the epitope depends on the strength and degree of this binding.<sup>202</sup>

The employment of isolated epitopes rather than the whole antigenic protein in vaccine formulations increases antibody production efficacy and, at the same time, provides an economic and synthetic advantage. Moreover, synthetic peptide epitopes have the advantage of easy characterization for purity and composition. A series of widely used T-cell epitopes in vaccine formulations derive from the diphtheria toxin (DT) or from the cross-reactive material 197 (CRM197), a non-toxic mutant of the diphtheria toxin.<sup>203</sup> CRM197 is antigenically similar to the native protein and enhances antibody response. A single mutation on the primary sequence of DT makes CRM197 harmless, yet immunologically indistinguishable from the native toxin.<sup>204</sup> Therefore, epitopes derived from DT have been widely used in vaccine production.<sup>205-207</sup> For the present study, we selected a T-cell epitope from DT that has the following sequence: H<sub>2</sub>N- QVVHNSYNRPAYSPG-COOH.<sup>208-210</sup> The epitope, here named Ep01, represents amino acids 379-383 of the DT sequence (535 residues in length in total).

### 5.1.3 The carrier molecule

The use of a carrier molecule for delivering both antigenic carbohydrate and T-cell epitope is an effective method for improving vaccine efficacy. The high surface-area-to-volume ratio and chemical diversity of nanoscale structures renders them of great interest to the biomedical and pharmaceutical fields. Inspired by molecular recognition processes in nature, a host of synthetic nanoscaffolds, including branched polymers and dendrimers, PNA, LNA and self-assembling peptides<sup>29, 31, 211, 212</sup> have been developed for numerous biomedical applications. However, the most widely used carriers for vaccines production are proteins. The primary carrier proteins currently available are the tetanus and diphtheria toxoids, and the derivative CRM197 of the latter. They are used precisely because they contain in their sequence epitopes recognized by the immune system, so that conjugation to the epitope is not necessary.

However, numerous concerns exist for the use of these proteins in vaccination. For example, the use of a limited number of suitable carrier proteins means that multiple vaccines will be built upon them. In the case of multiple vaccinations with material containing the same carrier protein, the probability of undesirable reactions is increased, since the presence of already existing antibodies could induce adverse systemic immunologic sensitivity reactions.

Moreover, as described by Schutze, phenomena of epitopic suppression can occur.<sup>213</sup> This event is observed when immunity to a protein contained in the conjugate already present in the vaccine interferes with the generation of a response to the covalently coupled antigen.

Finally, there are obvious increased costs in the production and purification of a full-length protein.

Clearly, there is an urgent need for alternative molecule as vaccine carriers that would overcome the immunological consequences of the repetitive use of the same scaffold protein and yet that would not interfere with the immunogenicity. For this reason, in this study we propose a self-assembling, fiber-forming peptide as a scaffold, here named FF03, for the presentation of a T-cell epitope and a carbohydrate antigen. This type of scaffold would offer synthetic simplicity, regular ligand distribution and, most importantly, high local antigen and epitope density.

It is now commonly accepted that high epitope density contributes to enhancing immunogenicity: T-cell epitopes with a higher concentration of MHC binding motifs are more immunogenic, probably due to the increased potential for being effectively processed by antigen-presenting cells. If more than one binding motif is present, a range of different MHC molecules can be activated and, as a consequence, a higher number of T-cell helper clones can



be generated, leading to a synergistic effect. We believe that a self-assembling peptide scaffold such as FF03 would enable molecular recognition by a synergistic combination of several cooperative binding events.

#### **5.1.4 The conjugation method**

The coupling of polysaccharide antigens or glycoconjugates to a peptide or a protein requires chemical activation of one or both components. The activation site on the glycan varies according to its structure; the main reactive group in proteins and peptides is either the amine at the N-terminus or the  $\epsilon$ -amine of a lysine residue (if present in the sequence). Random activation of the reaction partners would produce glycoconjugates with poorly defined structures. To gain control over the conjugation site, a linker arm between the carbohydrate and the protein/peptide must be introduced. Such spacers should be long enough to avoid steric clashes between the protein and the glycan but short enough to enable coupling; it should also contain a reactive functional group for facile and selective conjugation but should not interfere with the immunogenicity of the reaction partner. In the case of small oligosaccharides, like the one studied here, *Leishmania* LPG-derivative Man-Gal disaccharide, coupling to the protein carrier or to peptide epitopes can proceed via a specific synthetic route to yield well-defined conjugates in a reproducible way.<sup>214-216</sup>

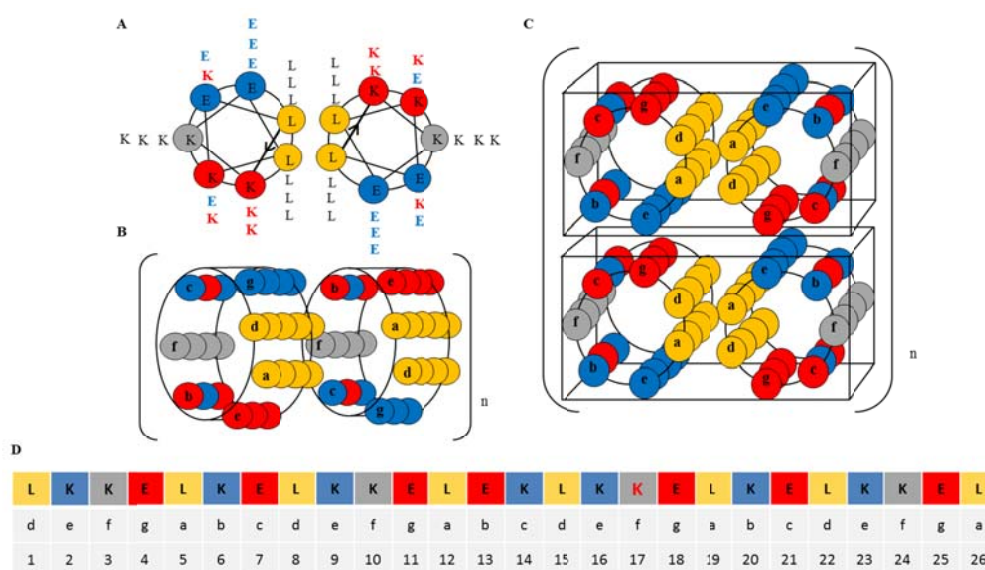
For this study, we used the same linker arm used by our collaborator at the MPI, presented in Figure 38B, a 5-hydrocarbon chain terminating with an amino group. This amine was used for the coupling of Man-Gal to the peptide epitope Ep01. Furthermore, the glycoconjugate antigen-epitope was coupled to the self-assembling peptide scaffold FF03 by means of the side chain of a selected lysine residue. Both conjugations were performed on solid phase, and the details are described in sections 5.2 and 5.3.

## **5.2 MULTIVALENT SCAFFOLD DESIGN**

As discussed in Chapter 1, multivalent interactions play an essential role in many biological processes. Multivalency represents one of nature's most ingenious methods to conduct and control the high number of recognition events necessary for cell survival, since it leads to an additive or exponential increase in binding affinity. For example, certain immune system receptors have evolved to distinguish among binding partners based not only the composition of the ligand array<sup>217</sup> but also on its density.<sup>218</sup> Glycans and peptide epitopes are useful

diagnostic probes in bioassays and immunogens in vaccines,<sup>219, 220</sup> and, due to their central role in the immune response, represent interesting classes of ligands to be explored in the context of self-assembling scaffolds.

With these criteria in mind, we designed a peptide nanoscaffold based on the coiled-coil folding motif that self-assembles into  $\alpha$ -helical fibers under physiological conditions. The peptide scaffold described here, FF03, was originally identified in the Ph.D. thesis of Dr. Enrico Brandenburg from 2012;<sup>221</sup> however, it was previously neither analyzed, nor employed in any application. FF03 contains three coiled-coil heptad repeats that constitute the core of the peptide sequence. An extra hydrophobic residue, leucine, is located at the C-terminus and half a heptad (*d, e, f* and *g*) at the N-terminus, to create the “sticky ends” effect that facilitates longitudinal assembly of the fibers along the helix axis (see section 1.2.2). Furthermore, fiber elongation is promoted by “head-to-tail” stacking of the helices. In this conformation, complementary charges easily interact with each other. Within each heptad, *a* and *d* positions are occupied by leucine residues, providing the hydrophobic core of the structure. To guarantee the formation of self-assembling fibers, lysine is present at all *e* positions and glutamate at all *g* positions, while the *b* and *c* positions carry these two charged amino acids in an alternating and complementary fashion. Thus, when the first heptad of the sequence contains lysine at its *b* position and glutamate at its *c* position, the second heptad in the sequence will carry glutamate at its *b* position and lysine at its *c* position. This design aims to induce the helices to interact with each other as well as to stabilize the coiled-coil motif. Figure 39 depicts a model for how fibers of FF03 could form.



**Figure 39.** FF03 fiber formation at neutral pH. Cylinders represent single helices; circles within each cylinder represent the 26 residues of the peptide sequence. A: helical wheel

representation of FF03; B: sticky end effect; C: electrostatic interaction contribution to fiber formation (boxes indicate dimers that share common hydrophobic interface); D: FF03 primary sequence and relative position within the heptad. Yellow: hydrophobic residues; red: positively charged residues; blue: negatively charged residues; grey: most solvent-exposed residues. Letters within circles give the single-letter code for amino acids (A) or indicate heptad repeat positions (B and C).

Electrostatic interactions between different dimers (shown boxed in Figure 39) are likely the main factor supporting growth of the fibers perpendicular to the helix axis. In fact, the *e* and *g* positions stabilize the core coiled-coil motif and the oligomeric state, while the *b* and *c* positions likely play a role in bundle formation, enabling association between fibers. As described in section 1.2.2, the presence of leucine residues at positions *a* and *d* typically leads to the formation of a trimer. Nevertheless, the peptide described in section 4.2 was shown to adopt a dimeric oligomerization state despite bearing a leucine zipper motif. The peptide scaffold FF03 has been designed following the same rules with the exception of the alternating charges on *b* and *c* positions. In an effort to rationalize fiber formation of FF03, we suggest the tetrameric arrangement shown in Figure 39, panel c.

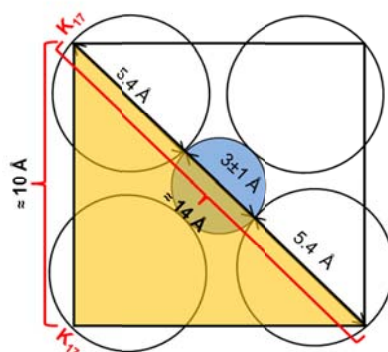
In all heptads, the *f* position, the most solvent-exposed position, is occupied by lysine residues. This amino acid was chosen for its side chain functionality, i.e., the side chain of the lysine on position 17 of the sequence (in red in Figure 39D), was selected to bear the antigen ligands.

### 5.3 GLYCOCONJUGATE SYNTHETIC STRATEGY

In order to design a glycoconjugate that would theoretically enable optimal display of the ligands to the appropriate binding partners, it must be taken into account that the immune response to carbohydrate-protein/peptide conjugates depends on: 1. the size of the carbohydrate; 2. the nature of the carrier molecule; 3. the nature and number of bonds between antigen moieties and the carrier; 4. the nature of the linker; 5. and the carbohydrate-to-carrier ratio.<sup>222</sup> We decided to employ a single, well-defined conjugation site on our peptide scaffold, the side chain of lysine 17, which occupies an *f* position in the heptad repeat. The *f* position of the heptad repeat is solvent-exposed, and we expected that ligands presented at this position would be readily available for antibody binding. Moreover, as discussed in section 5.1.3, the antigenic activity of the glycoconjugates generally depends on the density of

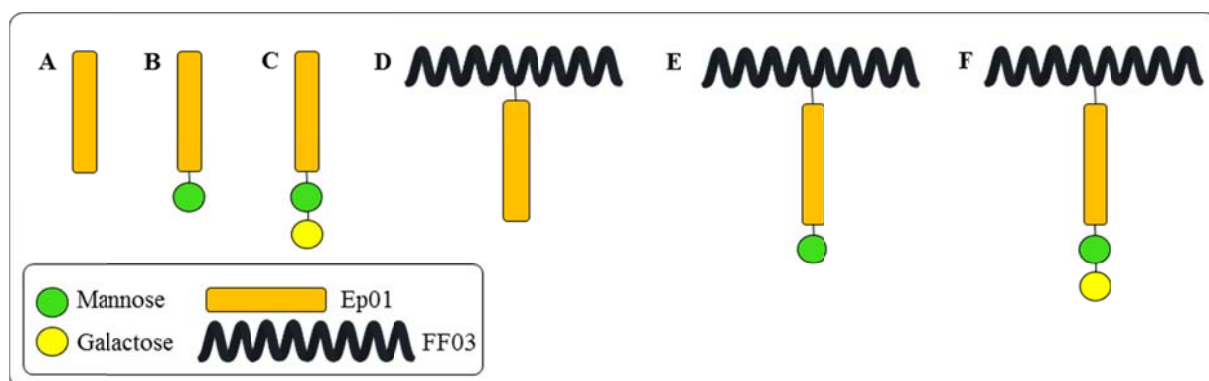
the carbohydrate on the carrier. Therefore, the large number of accessible antigens presented on our multivalent peptide scaffold would likely ensure a higher level of immunogenicity. If we assume a tight head-to-tail connection among the single peptide chains due to the sticky end effect, we can calculate that the ligands are presented on FF03 every 40 Å along the fiber's main axis (there are 10 Å between the *f* positions within one sequence and four full heptads between two lysine 17 residues bearing the ligands). However, the distance between the ligands displayed on the surface of proximal helices within a fiber (width of the single fiber) depends on the minimal oligomerization state of the peptide and on the number of fibers composing each bundle.

To calculate the nearest distance between two neighboring lysine 17 residues perpendicular to the fiber axis, we use as the starting point the fiber-formation model presented on Figure 39. In the model, the minimal oligomerization state for FF03 is tetrameric. The reported average diameter for the core of a tetrameric coiled-coil peptide is  $3 \pm 1$  Å.<sup>223</sup> By adding to this the measure of twice the width of an average helix ( $2 \times 5.4$  Å = 10.8 Å), we obtain an approximation of the distance between the two most solvent-exposed position residues of two helices placed opposite each other in a tetramer. This distance is close to 14 Å, and it represents the diagonal of the square inscribed in the circle described by the four helices. By dividing this square in two, we obtain two identical right triangles and we can therefore use the Pythagorean theorem to calculate the two shorter sides of it. This length represents the distances between the most solvent-exposed positions among each helix of the tetramer and, therefore, the approximate distance at which the ligands attached to this position are placed. This value is  $\approx \sqrt{\frac{14^2}{2}}$  Å  $\approx 10$  Å (Figure 40).



**Figure 40.** Approximate calculation of the distance between two ligand-bearing amino acids within proximal helices. Empty circles: top view of coiled-coil constituting the tetramer; blue circle: diameter of the interhelical space based on crystal data; yellow triangles: right triangle selected to apply the Pythagorean theorem.

To verify whether the self-assembling peptide scaffold FF03 could be employed for the multivalent presentation of antigens for vaccination and diagnostic purposes, we created a small library for the display of the T-cell epitope Ep01, and the carbohydrate antigen Man-Gal. As proof of concept, also the monosaccharide mannose (Man) was included as ligand. Both Man-Gal and Man were modified with the aminopentanyl linker shown in Figure 38. A cartoon of the peptide library is depicted in Figure 41, and details are provided in Table 5.



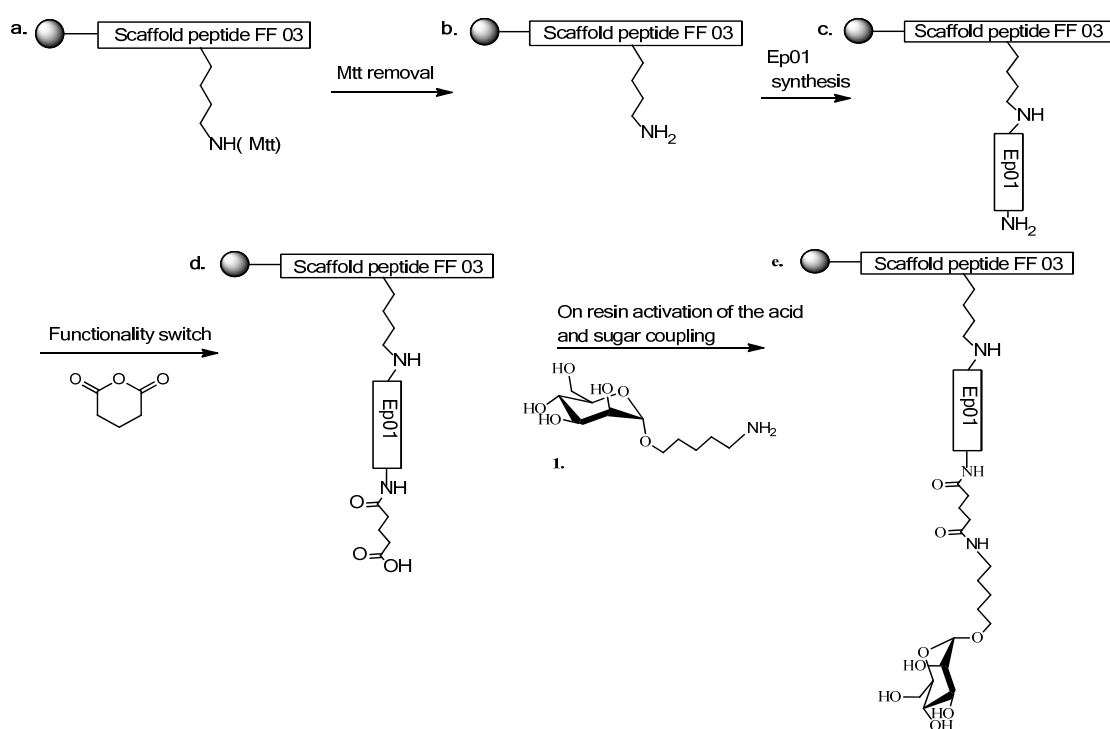
**Figure 41.** Peptide library. A: Ep01; B: Ep01-Man; C: Ep01-Man-Gal; D: FF03-Ep01; E: FF03-Ep01-Man; F: FF03-Ep01-Man-Gal.

NAME		SEQUENCE AND DESCRIPTION
A	Ep01	DT epitope H <sub>2</sub> N-QVVHNSYNRPAYSPG-OH
B	Ep01-Man	DT epitope with 5-amino-pentanyl- $\alpha$ -D-mannopyranoside
C	Ep01-Man-Gal	DT epitope with leishmanial lipophosphoglycan
D	FF03-Ep01	Scaffold H <sub>2</sub> N-LKKELKELKKELEKLK <b>K</b> ELKELKEL-OH with DT epitope
E	FF03-Ep01-Man	FF03-Ep01 with 5-amino-pentanyl- $\alpha$ -D-mannopyranoside
F	FF03-Ep01-Man-Gal	FF03-Ep01 with leishmanial lipophosphoglycan

**Table 5.** Peptide sequences and description of the peptide library. The particular lysine to which the different ligands are attached is shown in bold.

As for the glycopeptides in the study described in chapter 4, the synthesis of the whole library was performed on solid phase. For the peptides without scaffold (Figure 41A-C), subsequent to synthesis of the peptide epitope Ep01, the functionality at the N-terminus was switched from an amine to a carboxylic acid by means of glutaric anhydride. After activation of the acid on resin, the carbohydrates Man and Man-Gal were conjugated via the aminopentanyl linker arm present on the hydroxyl group on their anomeric carbons, to generate the glycopeptides Ep01-Man and Ep01-Man-Gal, respectively.

In the case of the components of the library containing the peptide scaffold FF03, the N-terminus of FF03 was blocked with aminobenzoic acid (Abz). After selective deprotection of the lysine at position 17, bearing the 4-Methyltrityl (Mtt) protecting group, removable under mild acidic conditions, the free amino group of the side chain was used to build up the peptide ligand Ep01. After switching functionality of the Ep01 N-terminus by means of glutaric anhydride, the resulting free carboxylic acid was activated while still “on resin” and the sugar moiety coupled via a short amino-linker. This strategy allows for the synthesis of glycopeptide conjugates entirely on solid phase. A scheme of the solid-phase synthesis strategy adopted to build up the glycopeptide library is given in Figure 42.



**Figure 42.** Scheme of the synthetic strategy adopted for the glycopeptide FF03-Ep01-Man.

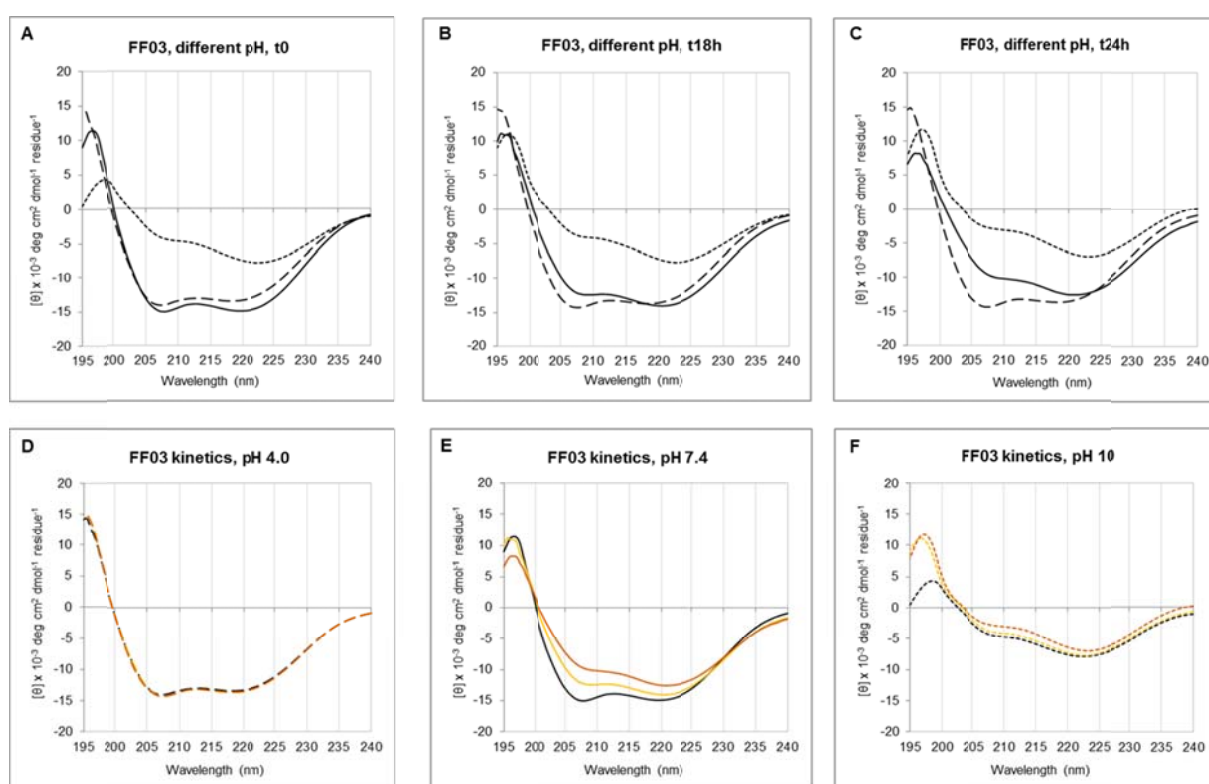
#### 5.4 Evaluation of peptide library structural features

The secondary and quaternary structure of the peptide scaffold plays a fundamental role in the realization of the desired multivalent effect. FF03 was evaluated for its suitability to serve as a stable scaffold for the multivalent presentation of the chosen glycopeptide ligand by self-assembling into a defined three-dimensional architecture. Also the structure of Ep01 was explored, since the epitope had not yet been structurally characterized outside the context of the diphtheria toxin.

The secondary structure of all library members was investigated by circular dichroism (CD), while the presence of fibers was detected by transmission electron microscopy (TEM).

### 5.4.1 Peptide-library secondary structure

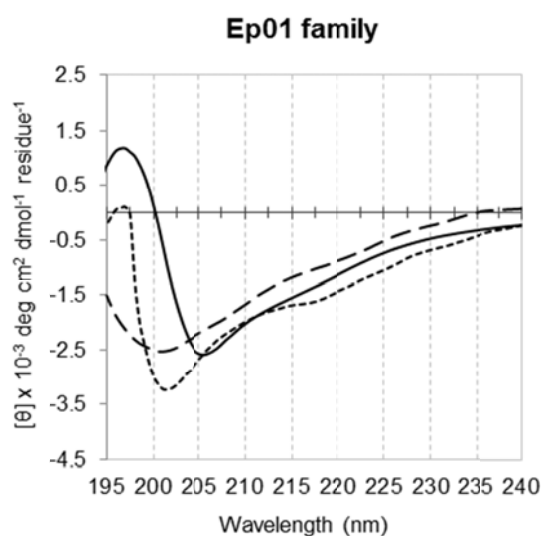
The peptides' spectroscopic analysis, as well as all the biological assays that followed, were performed under physiologically relevant conditions. Nevertheless, to gain a better understanding of the fiber-formation process, the secondary structure of FF03 was analyzed via CD at different pH values and time points. The results of these investigations are reported in Figure 43.



**Figure 43.** FF03 CD spectra collected at concentrations of 150  $\mu$ M, at different pH and time from sample preparation. Solid line: FF03 at pH 7.4; short dashes: FF03 at pH 10; long dashes: FF03 at pH 4.0. A-C: FF03 secondary structure comparison at three different pH values immediately (A), after 18 hours (B) and after 24 hours (C) from sample preparation. D-F: FF03 secondary structure comparison at different time points at the pH values 4.0 (D), 7.4 (E) and 10 (F); black lines indicate t0, yellow lines indicate t18h and orange lines indicate t24h.



CD spectra acquired at different pH values show different degrees of ellipticity, but always describe a CD profile typical of a purely  $\alpha$ -helical peptide, with minima at 208 nm and 222 nm and a maximum at 195 nm. At physiological pH and a concentration of 150  $\mu$ M, the intensity of the minimum at 208 nm gradually decreases over 24 hours, indicating a tendency to form higher supramolecular structures.<sup>224</sup> CD spectra recorded at pH 4.0 and 10 show different behaviors: at pH 4.0 no change on the spectrum was detected over the 24-hour time course, indicating that, under these conditions, only smaller oligomers are formed; at pH 10, FF03 is mostly insoluble (white precipitate is observed during sample preparation), but the aggregates that form show similar CD traces but lower degrees of ellipticity. This could indicate the rapid formation of larger bundles, which may become insoluble due to their size. The CD data of the other components of the peptide library are reported in Figures 44 and 45.

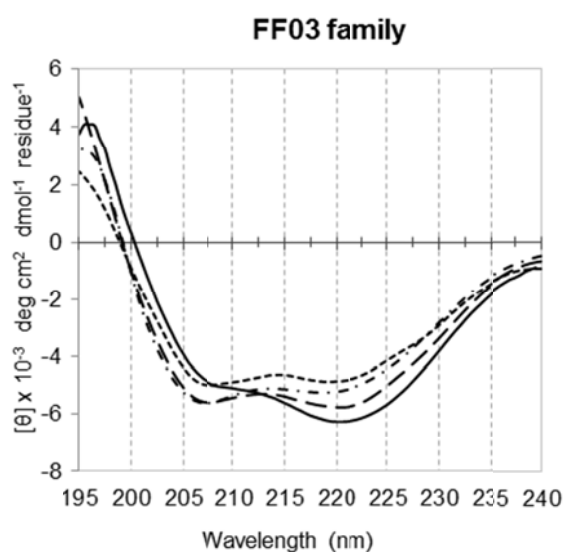


**Figure 44.** CD spectra of the Ep01 family collected 24 h after sample preparation at 100  $\mu$ M. Solid line: Ep01; short dashes: Ep01-Man; long dashes: Ep01-Man-Gal.

Ep01, which, within the DT structure, is partially folded into an  $\alpha$ -helix,<sup>225</sup> when isolated appears to adopt a combination of secondary structures (Figure 44). The global minimum between 200 and 205 nm suggests either a combination of random coil and  $\beta$ -sheet conformations, or the presence of  $\beta$ -turns,<sup>226, 227</sup> while the maximum at 195 nm could indicate the co-existence of  $\alpha$ -helical species. The spectrum of Ep01-Man differs from that of the parental peptide only in a shift of the minimum at 205 nm towards shorter wavelengths, whereas Ep01-Man-Gal displays no maximum at 195 nm but an otherwise similar tendency to the other two traces. Ep01 was used exclusively at low concentrations and within 24 hours after sample preparation in all experiments below.



The FF03 derivatives show a CD profile typical of an  $\alpha$ -helical secondary structure (Figure 45). There are, however, subtle differences between unmodified FF03 and the three modified species. As discussed, the parental peptide FF03 displays a less intense minimum at 208 nm wavelength, likely attributable to helical fiber formation.<sup>228</sup> FF03-Ep01, FF03-Ep01-Man and FF03-Ep01-Man-Gal display a shorter-wavelength minimum with increased intensity at 208 nm, resulting in roughly equal intensities for the two minima at 208 nm and 222 nm. This difference in CD behavior could be due to either reduced fiber formation in the modified FF03 species, or to the presence of Ep01, which has a minimum between 200 nm and 205 nm.



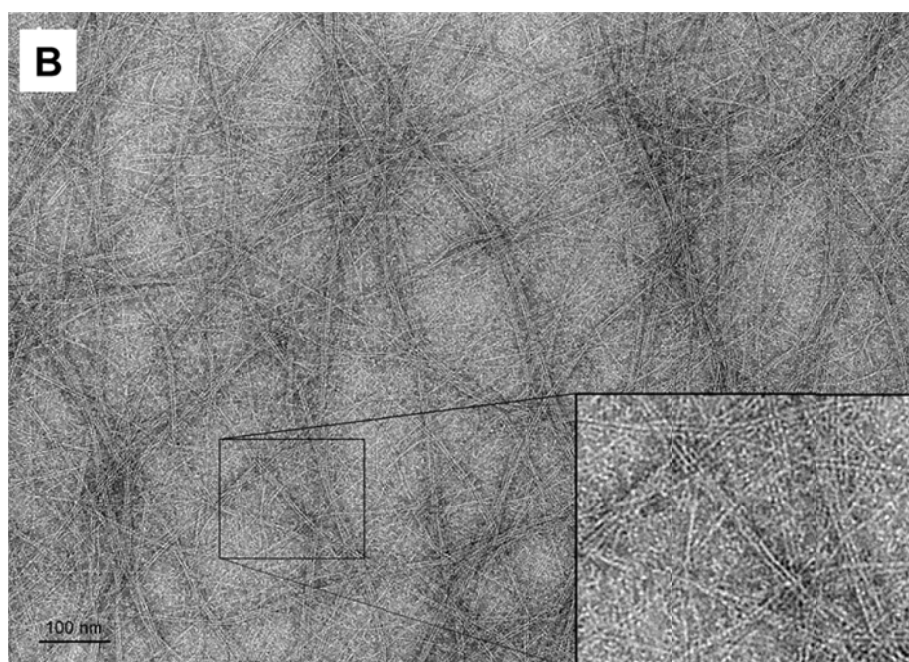
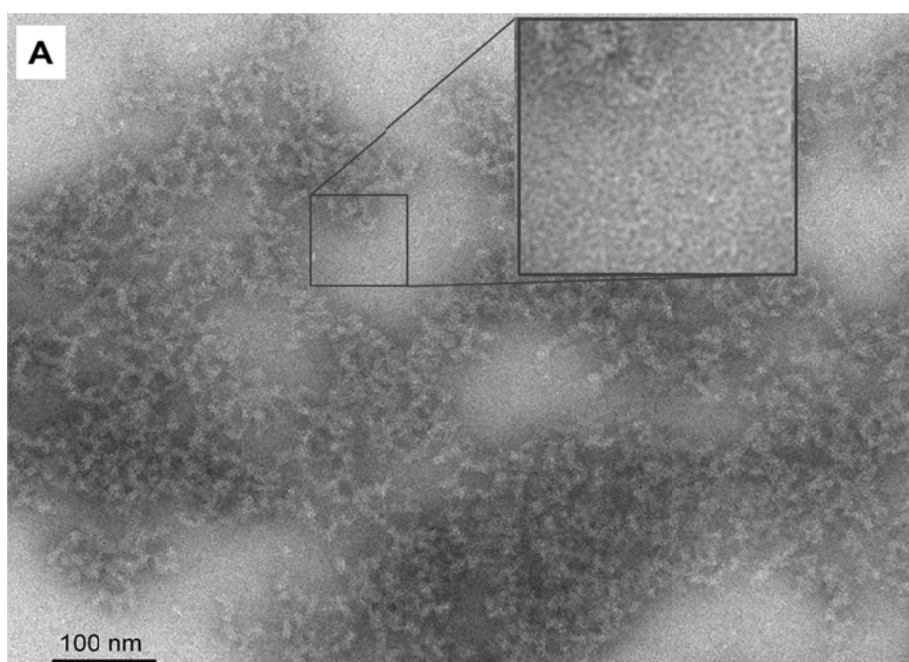
**Figure 45.** CD spectra of the FF03 family collected 24 h after sample preparation at 100  $\mu$ M. Solid line: FF03; short-long dashes: FF03-Ep01; short dashes: FF03-Ep01-Man; long dashes: FF03-Ep01-Man-Gal.

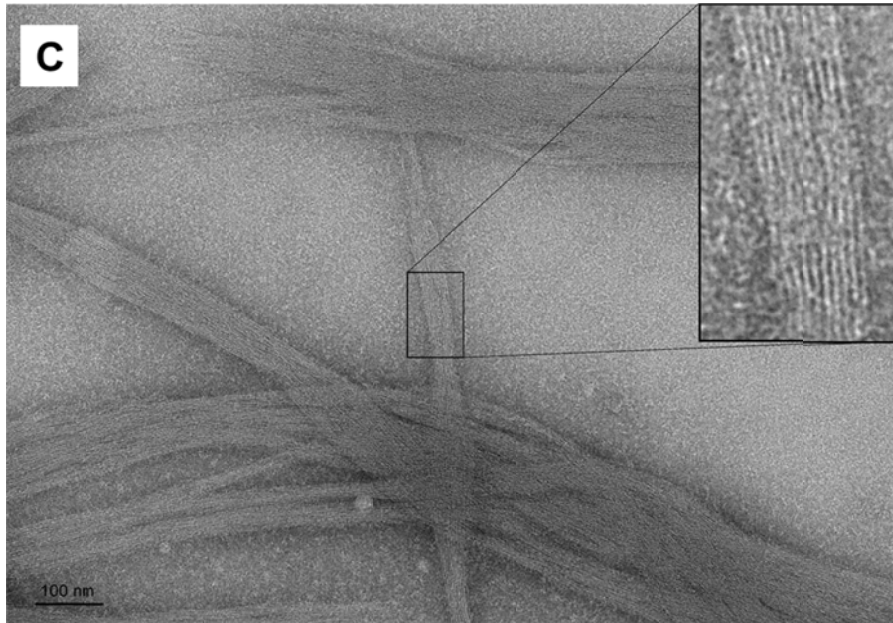
The results of the CD analysis suggest that the structure of the peptide scaffold FF03 is robust enough to retain its  $\alpha$ -helical character even after heavy modification with both the T-cell epitope Ep01 and the carbohydrate antigen Man-Gal. Stability in the secondary structure evidenced by a comparable degree of ellipticity in the CD spectra guarantee regular distribution of the epitope and antigen ligands along the helices. Nevertheless, the formation of the fibers is crucial for reaching the desired multivalent effect; therefore their presence was verified for all glycoconjugates via TEM.

### 5.4.2 Peptide library supramolecular structure

TEM analysis was carried out to determine whether FF03 and its derivatives are in fact able to self-assemble into  $\alpha$ -helical fibers, as suggested by CD analysis. As controls, Ep01 and its glycosylated versions Ep01-Man and Ep01-Man-Gal were also analyzed; as expected, no aggregates were found in these three cases.

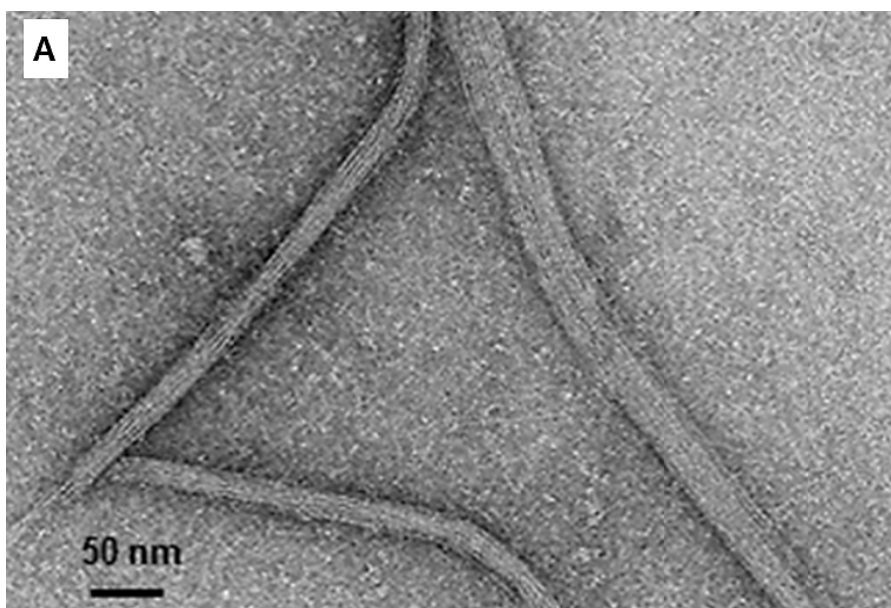
The peptide scaffold FF03 was analyzed within one hour from sample preparation at the pH values 4.0, 7.4 and 10.0 and after 24 hours and one week from sample preparation at pH 7.4. This study would help understanding the kinetics of fiber formation and their stability in solution. The results of these investigations are reported in Figures 46 and 47.



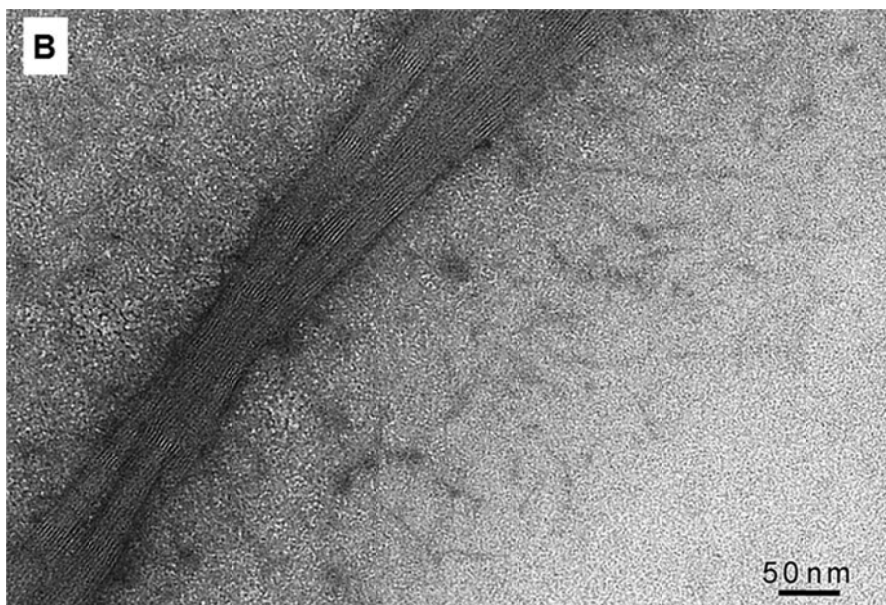


**Figure 46.** TEM micrographs of FF03 acquired at a concentration of 100  $\mu\text{M}$ , 1 h from sample preparation at pH 4.0 (A), pH 7.4 (B) and pH 10 (C). Square: “zoom in” of indicated section of micrograph.

The peptide scaffold FF03 indeed forms  $\alpha$ -helical fibers, and neutral conditions are preferred for obtaining soluble regular supramolecular structure. At pH 7.4, FF03 forms long fibers and bundles, and its viscosity is similar to that of a hydrogel. TEM micrographs of FF03 at pH 4.0 show no long fibers, but only flocculent precipitates. Basic pH leads FF03 to assemble into very thick bundles and does not allow for complete dissolution of the peptide.



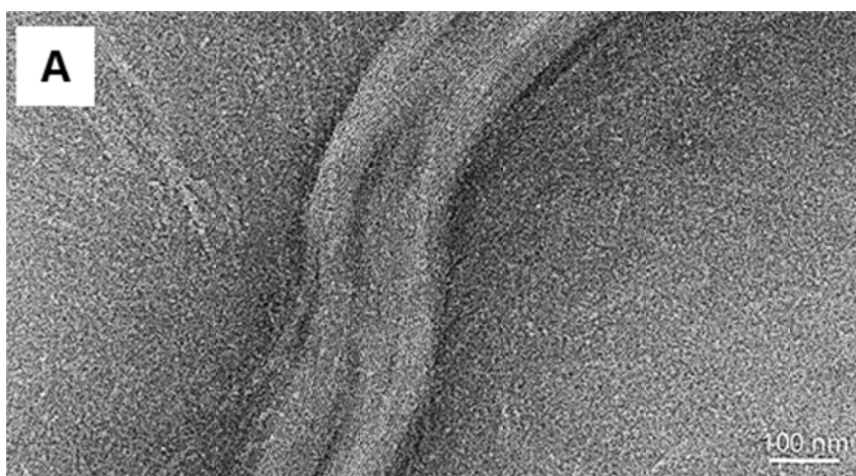


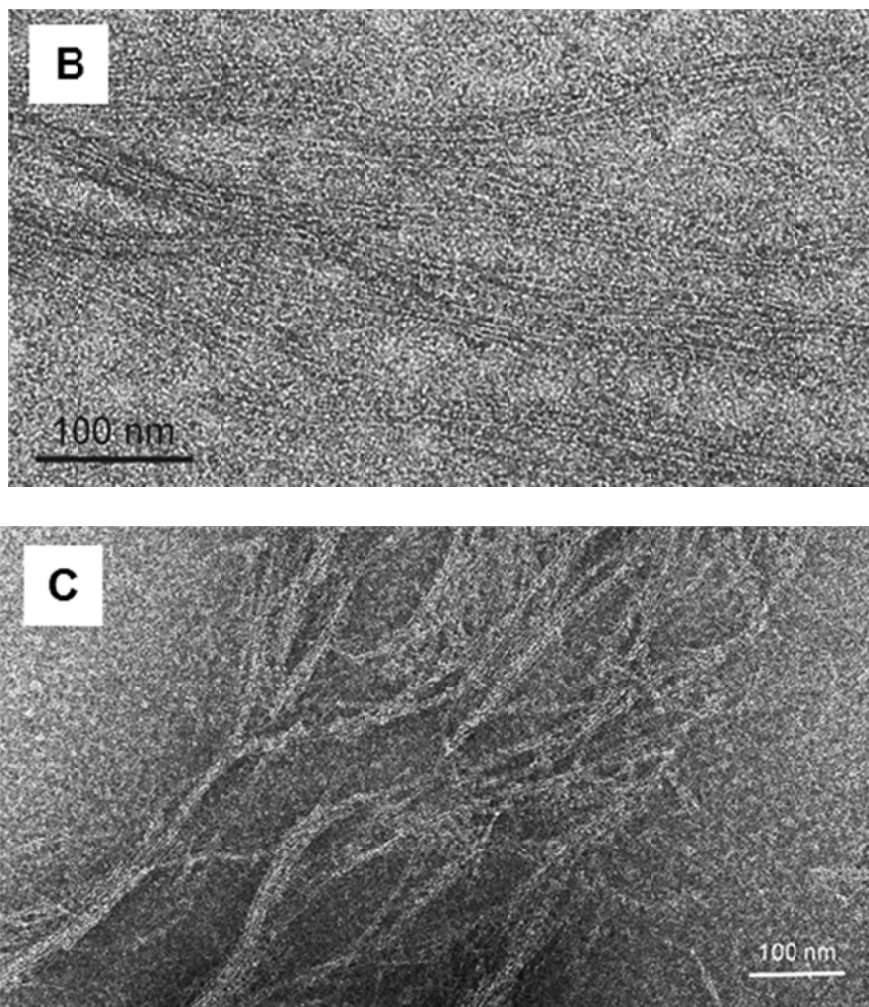


**Figure 47.** TEM micrographs of FF03 at pH 7.4 acquired at a concentration of 100  $\mu$ M and at A: 24 h from sample preparation; B: one week from sample preparation.

The fibers formed by FF03 are exclusively linear and non-branching. The self-assembly process likely begins immediately after dissolving the peptide in neutral buffer and is complete after 24 hours: TEM micrographs acquired seven days from sample preparation show fibers of comparable type, length and thickness to the ones observed after only 24 hours. The formation of bundles appears to be a signature of the FF03 series, and indicates strong interfiber interactions.<sup>90</sup> The isolation of a single fiber is therefore often not feasible.

In order to determine whether peptide scaffold FF03 is able to retain its ability to self-assemble when loaded with the Ep01 epitope and the carbohydrates Man or Man-Gal, additional TEM experiments were carried out. The results are reported in Figure 48.





**Figure 48.** TEM micrographs of FF03 derivatives acquired at a concentration of 100  $\mu$ M 24 hours after sample preparation. A: FF03-Ep01; B: FF03-Ep01-Man; C: FF03-Ep01-Man-Gal.

The micrographs clearly demonstrate that all peptides from the FF03 series are able to form  $\alpha$ -helical fibers of a thickness of 2.5 nm thick and length of at least 100 nm long fibers; these individual fibers were observed alongside large bundles of fibers. Interestingly, it appears that the thickness of the bundles formed by the glycoconjugates (FF03-Ep01-Man and FF03-Ep01-Man-Gal) is reduced compared to FF03 and FF03-Ep01. This might be due to steric hindrance of the carbohydrate moieties placed, together with the peptide epitope, on the most solvent-exposed position of the scaffold, or to their interaction. Nevertheless, neither the presence of Ep01 nor its glycosylation influences the structure of FF03 enough to abrogate self-assembly or lead to the formation of other types of supramolecular structures, such as nanotubes or vesicles. This indicates that the fibers formed by the coiled-coil-based scaffold FF03 are stable enough to ensure the multivalent presentation of both peptide epitope and carbohydrate antigen.

## 5.5 MULTIVALENT PRESENTATION OF LIGANDS TO SELECTED BINDING PARTNERS

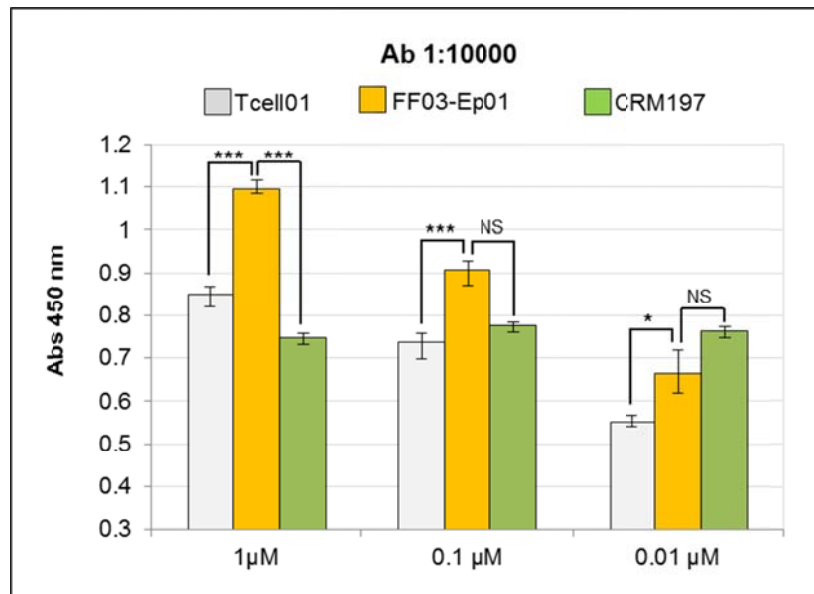
The multi-component strategy employed in this study ultimately aims to activate, simultaneously, different processes in the immune response. The main requirement is to cluster both T-cell epitopes and carbohydrate antigens to obtain higher-level antibody recognition. With the structural studies reported in section 5.4 we were able to demonstrate that multivalency can indeed be reached via our coiled-coil based peptide scaffold. The presence of a multivalent scaffold, though, does not imply that the displayed ligands are optimally accessible for binding with specific binding partners. To attempt to quantify the multivalent effect within this regime, we carried out a series of bioimmunological assays, described below.

### 5.5.1 Interaction of the peptide epitope with anti-diphtheria toxin antibodies

An enzyme-linked immunosorbent assay (ELISA) was used to determine how efficiently the monomeric Ep01 system and the multivalent FF03-Ep01 are bound by polyclonal antibodies against DT. The peptide Ep01, the conjugate FF03-Ep01 and the positive control CRM197 were immobilized on microtiter plates and incubated with sera from mice immunized with the CRM197.<sup>200</sup> Subsequent to primary antibody incubation, the plates were repeatedly washed with the chaotropic  $\text{NH}_4\text{SCN}$ , a dissociating agent, to reduce nonspecific binding.<sup>229</sup> Epitope recognition was determined by means of an anti-mouse IgG horseradish peroxidase-conjugate. The assay was performed using serial antibody dilutions (1:10000-1:1000000) and three peptide/protein concentrations (1 $\mu\text{M}$ , 100 nM and 10 nM). The results are reported in Table 6 and Figure 49.

Name	Serum dilution	Peptide/Protein concentration		
		1 $\mu\text{M}$	100 nM	10 nM
Ep01	1:10000	0.63473368	0.52190032	0.33683366
	1:100000	0.50783366	0.43930034	0.31681698
	1:1000000	0.45563366	0.36103366	0.28191702
FF03	1:10000	0.8165333	0.6272667	0.3833
	1:100000	0.6585334	0.4789667	0.29345
	1:1000000	0.4425	0.3761	0.28335
CRM197	1:10000	0.74853333	0.77526669	0.7642
	1:100000	0.70263335	0.73986666	0.74685
	1:1000000	0.69923334	0.72010001	0.67485

**Table 6.** ELISA results. Reported in the last column on the right are the average values of the absorbance at 450 nm, determined in triplicate.



**Figure 49.** Anti-diphtheria toxin antibody binding to the monomeric epitope (Ep01, grey columns), epitope bound to the multivalent scaffold (FF03-Ep01, yellow columns) or to CRM197 protein (green columns). Assays were performed with 1  $\mu\text{M}$ , 0.1  $\mu\text{M}$  and 0.01  $\mu\text{M}$  analyte and a 1:10000 serum dilution. Mean AUs at 450 nm (triplicate) are given. Asterisks indicate *P* values (\*\*\*: *P* < 0.005; \*: *P* < 0.05). NS = non-significant.

At an antibody dilution 1:10000 we observed the most significant differences among the data sets. It appears evident that anti-DT antibodies recognize the epitope Ep01 in all considered presentation forms (as monomer, presented on the multivalent scaffold FF03 and as part of the CRM197 whole protein). This means that the peptide scaffold FF03 is able to present the ligand epitope Ep01 to appropriate binding partners in a way that provides for recognition and interaction. The positive control CRM197 does not undergo reduction of antibody binding efficacy in a concentration-dependent fashion. The advantage here is evident: CRM197 can be used at the low nanomolar range and still be recognized and bound by specific antibodies. At the concentrations of 1  $\mu\text{M}$  and 0.1  $\mu\text{M}$ , the isolated epitope Ep01 shows immunogenicity comparable to CRM 197, but at the lowest analyzed concentration of 10 nM the antibody binding is reduced to about 65%. At this concentration, it is obviously more efficient, in terms of gained immunogenicity, to employ the whole CRM197 protein, rather than the isolated epitope.

The results relative to the antibody recognition of the epitope presented on the multivalent scaffold are very encouraging: at the higher concentration of 1  $\mu\text{M}$  we observed significantly higher absorbance values for FF03-Ep01 compared to Ep01 alone and CRM197, indicating

more efficient antibody binding (no binding towards FF03 alone was detected). Since the same amount of epitope was incubated with the 96-well plate in all cases, the improved antibody avidity is probably due to a more suitable organization of the epitope when presented on the fiber-forming scaffold. At lower concentrations, the difference in antibody binding between FF03-Ep01 and CRM197 is not statistically significant. Yet, the increased antibody binding of FF03-Ep01 compared to Ep01 retains its statistical significance at all tested concentrations.

The results do not only demonstrate that the multivalent presentation of Ep01 on FF03 is more efficient compared to the isolated epitope, but that, at higher concentrations, this coiled-coil scaffold is a better carrier protein than CRM197 itself, since it increases the immunogenicity of the presented epitope. We believe that the concentration-dependency observed for the FF03 conjugates is strictly correlated with the fiber formation process, which is a concentration- and time-dependent phenomenon.

Antigen platforms of the kind offered by peptide FF03 could find application not only in vaccine preparations, but they could also be used to enhance the sensitivity of diagnostic assays, regardless of whether high local concentration or the regularity of the spatial distribution of the ligand is the key factor in increased antibody recognition.

### **5.5.2 Interaction of the mannose ligand with lectins**

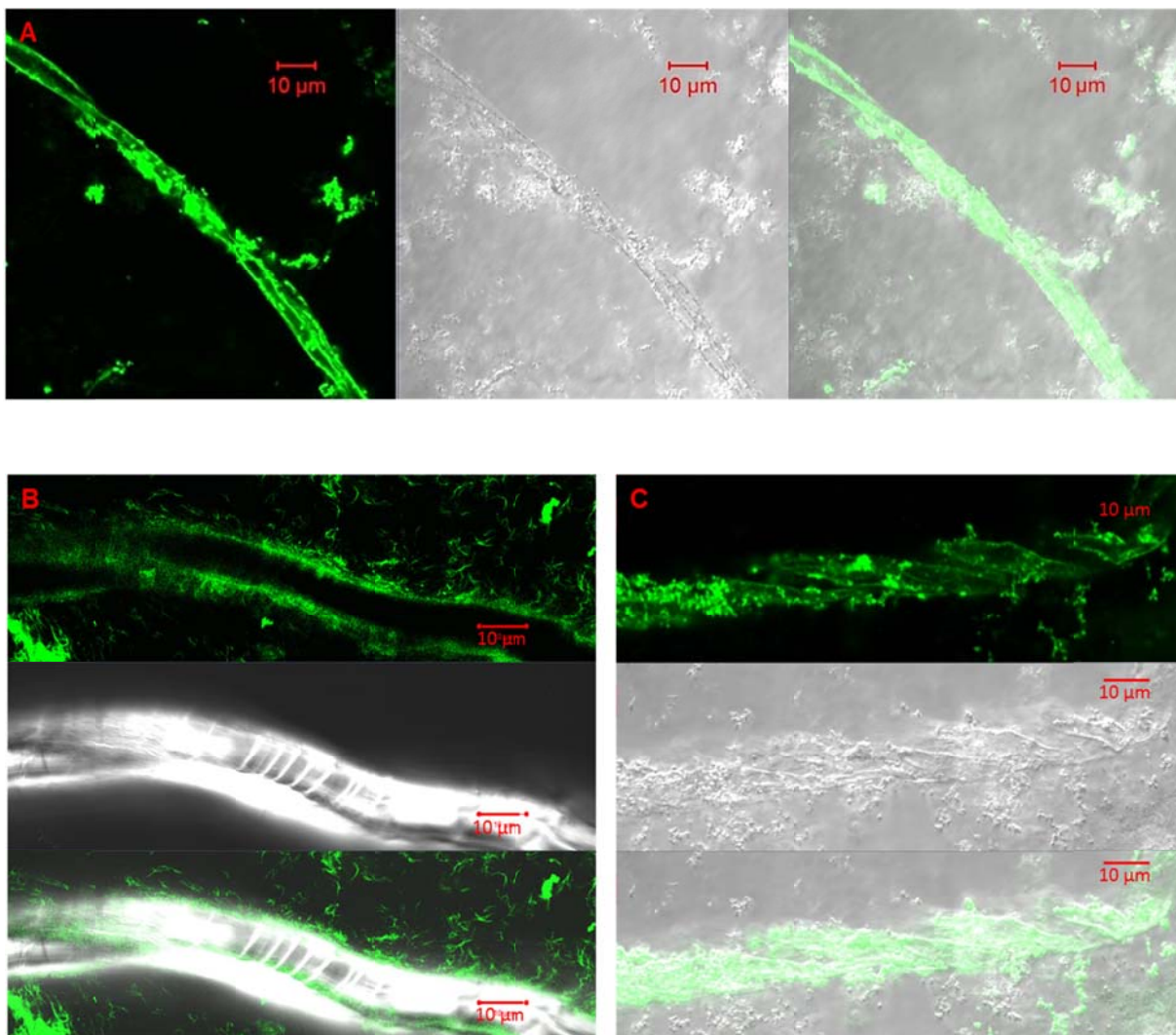
From the perspective of the host organism, unique glycan motifs serve as biomarkers for diseased cells and characterize invading pathogens. Lectins are naturally-occurring carbohydrate-binding proteins that are often employed to probe cell surface changes *in vitro*.<sup>230</sup> Due to the relatively weak individual intermolecular interactions that are formed between lectins and carbohydrates, these proteins have evolved to detect glycans that are organized in a multivalent fashion, and to discriminate among ligands based on glycan density.<sup>231</sup> For this reason, lectins are ideal binding partners to be employed to examine the effect of the multivalent scaffold FF03 on carbohydrate presentation. In order to verify the ability of the scaffold FF03 to efficiently present a carbohydrate ligand to selected lectins, we have employed the glycoconjugate FF03-Ep01-Man and the mannose-binding lectin ConA. Furthermore, we have tested the construct in a whole-cell assay by employing an *E. coli* strain expressing a pattern of surface Man-binding lectins. The results are reported below.

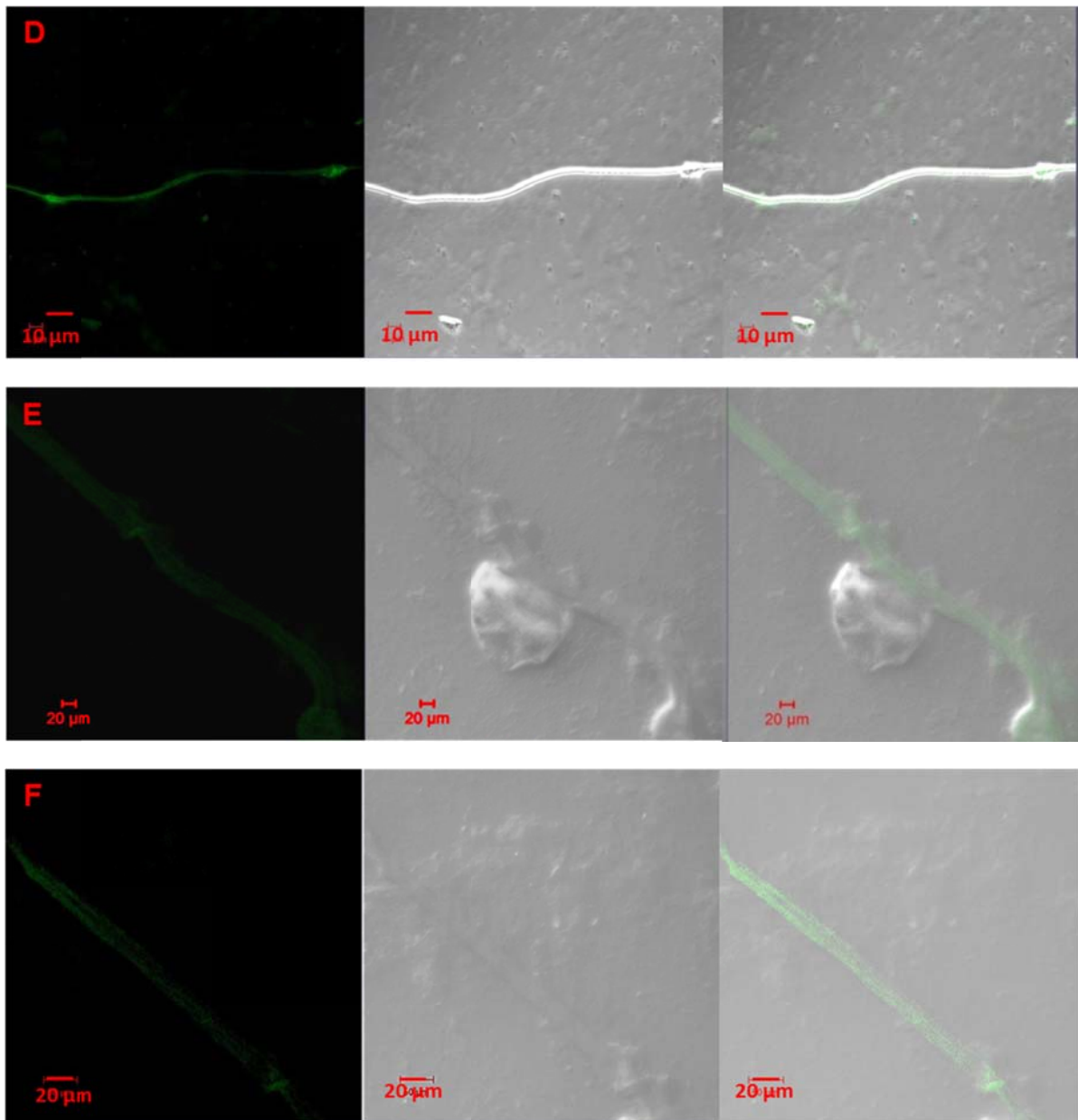


- **ConA-binding assay**

Concanavalin A (ConA) is a carbohydrate-binding protein (lectin) that specifically binds nonreducing terminal  $\alpha$ -D-mannosyl moieties. It is widely used in biology to characterize glycoproteins and glycoconjugates on cell surfaces. Like most lectins, it is naturally present as an oligomer (tetramer), with one mannose-binding domain per subunit.

To determine whether a fluorescently labelled lectin (ConA-FITC) is able to recognize the mannose moieties presented on the self-assembled FF03 scaffold, we used confocal laser microscopy (CLM). Green fluorescence relative to ConA-FITC must be compared in the presence or absence of mannose. Some representative CLM images are reported in Figure 50.



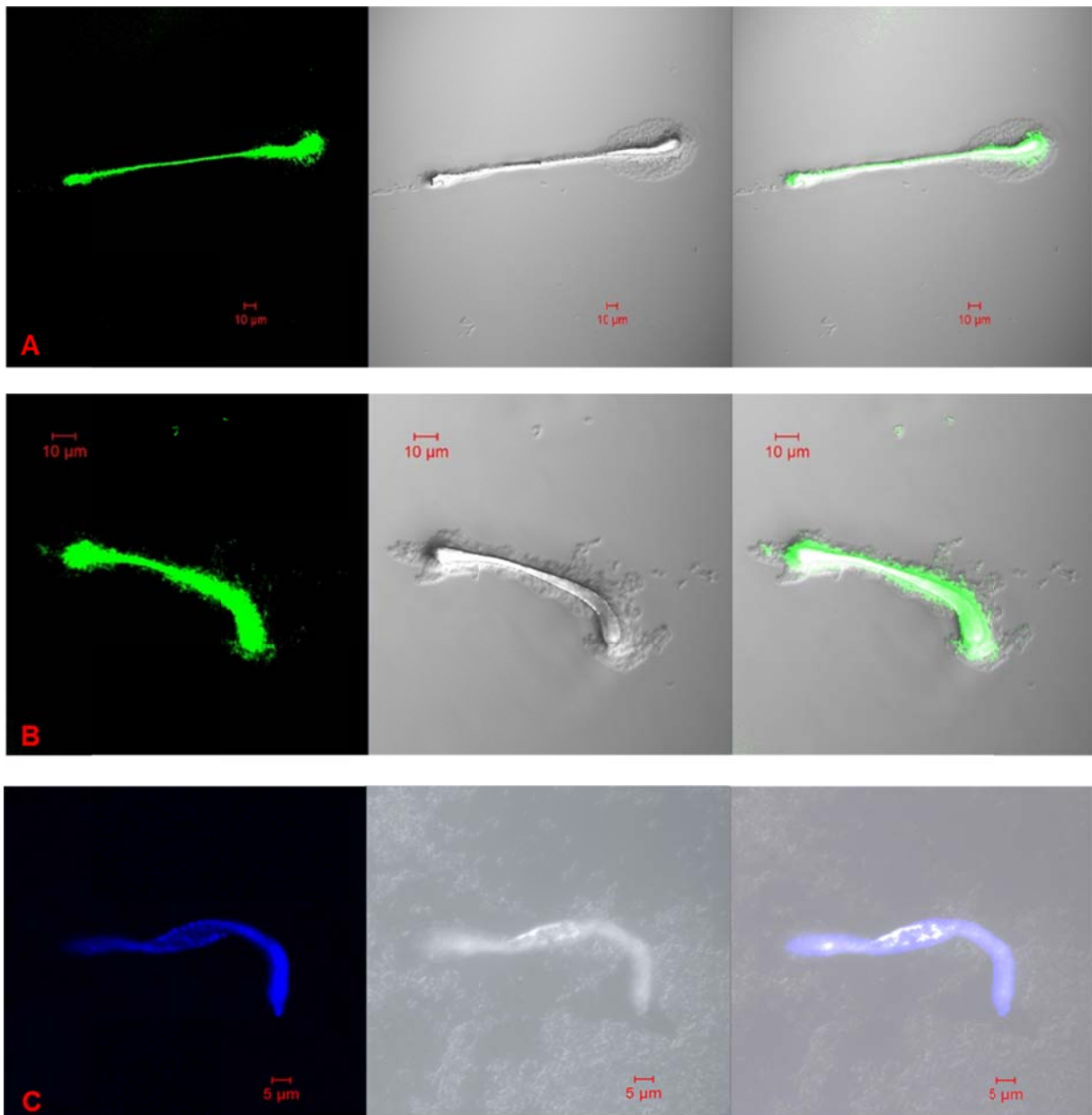


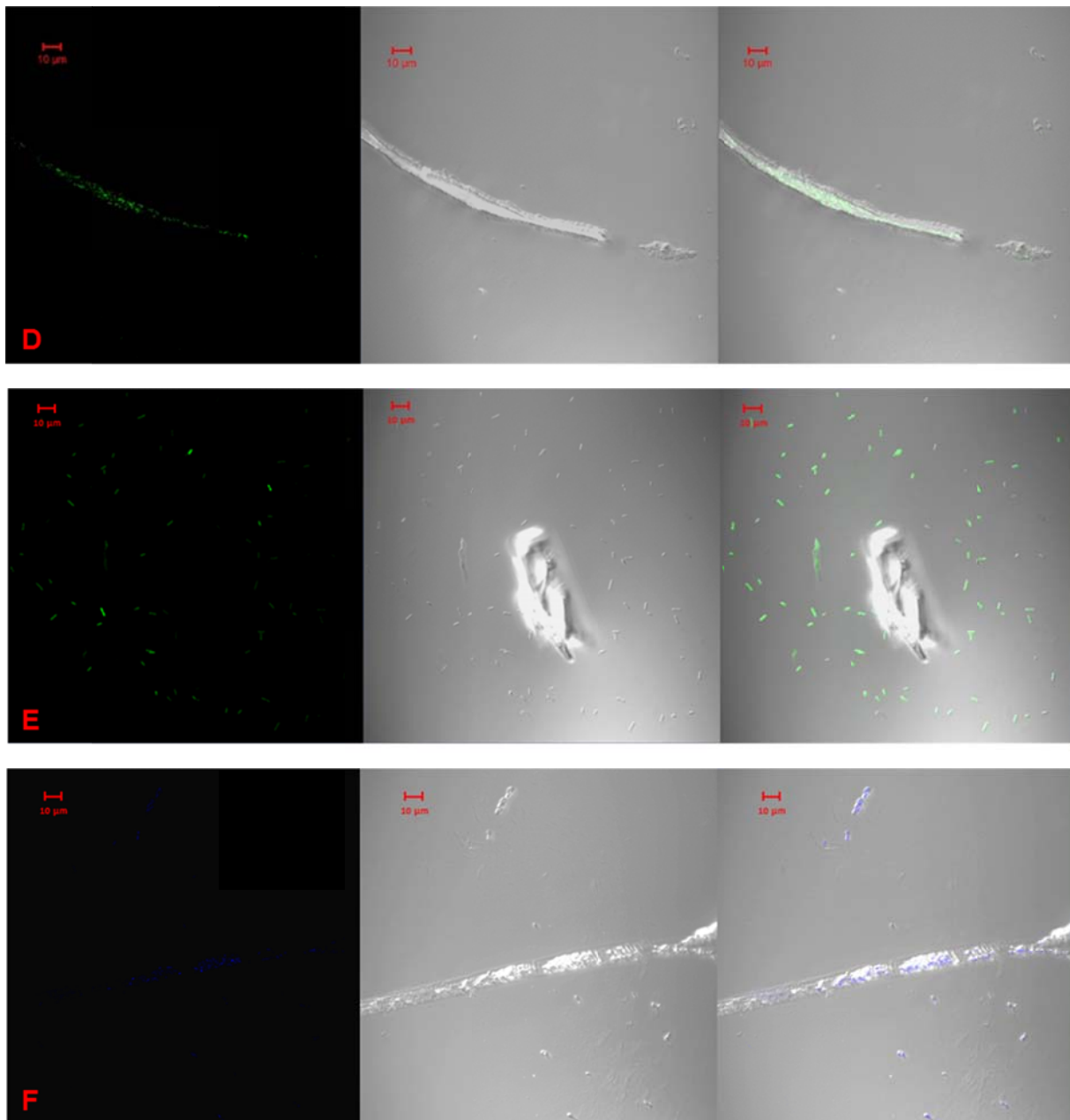
**Figure 50.** CLM images of the self-assembled scaffold after incubation with FITC-labeled ConA. Green fluorescence associated with FITC indicates efficient binding of ConA to the fiber-forming peptide. A-C: FF03-Ep0-Man; D-F: FF03-Ep01.

CLM images clearly reveal strong ConA fluorescence in the presence of mannosylated fibers (Figure 48A-C), while the images of the control peptide FF03-Ep01 show negligible fluorescence (Figure 48D-F). The fluorescence signal clearly defines the shape and organization of the macrostructures, suggesting that the carbohydrate moiety is highly available for lectin binding and that the coiled coil-based scaffold is an appropriate system for facilitating protein-glycan interactions.

- **Bacterial binding assay**

Ligand density and orientation are key factors for the success of *in vitro* diagnostic assays. To test whether our fiber-forming scaffold can compete with the standard carrier proteins currently used in diagnostic assays, we employed CLM to detect the interaction of FF03-Ep01 and FF03-Ep01-Man with *E. coli* ORN178, a strain expressing a certain pattern of lectins that enables specific binding to mannose. To visualize the results, ORN178 was conjugated with FITC (green fluorescence) and its nuclei were stained with DAPI (blue fluorescence). The results obtained in this whole-cell assay should be analogous to the ones observed for the ConA binding assay. Representative CLM images are reported in Figure 51.





**Figure 51.** CLM images of self-assembled scaffold after incubation with FITC-labeled and DAPI-stained *E. coli* ORN178. Green fluorescence associates with FITC; blue fluorescence associates with DAPI. Their intensity indicates efficient binding of *E. coli* to the fiber-forming peptide. A-C: FF03-Ep01-Man; D-F: FF03-Ep01.

Analogous to the results obtained from the ConA-binding assay, the fluorescent signal is much more intense and localized when mannose is present on the FF03-Ep01 scaffold, while faint, nonspecific binding of the labeled *E. coli* can be observed on the mannose-free FF03-Ep01. This finding confirms that the coiled-coil-based self-assembling system can efficiently present a monosaccharide ligand to a cell surface receptor and could therefore be employed in carbohydrate/lectins arrays.

### 5.5.3 Presentation of the antigenic disaccharide to specific antibodies

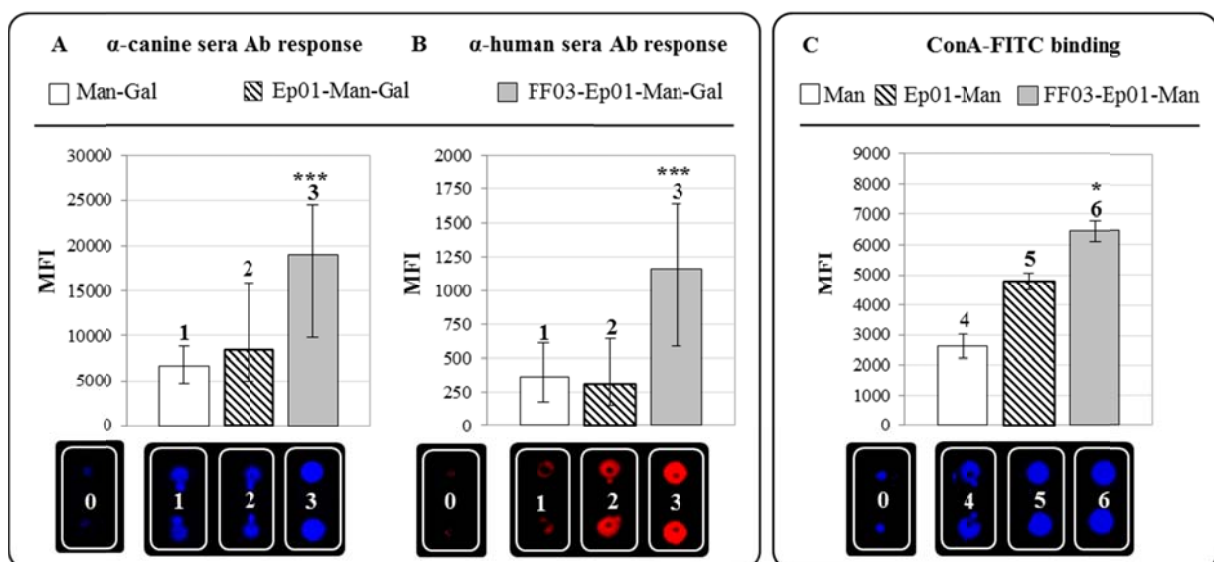
The ELISA, ConA and bacteria-binding experiments described above demonstrate that the coiled-coil-based self-assembling peptide is a suitable scaffold for the multivalent presentation of a peptide epitope and a monosaccharide. These results prompted further investigations to analyze the effective multivalent presentation of the antigenic disaccharide here called Man-Gal, the minimal epitope of a LPG found on pathogenic Leishmania parasites. The accessibility of the naturally-occurring carbohydrate antigen was evaluated by means of a microdot assay. Compared to the more conventional ELISA, the microdot assay allows for a high-number repetitions of the same conditions from one analyte sample, providing greater statistical significance. We employed a 96-grid glass plate, onto which we printed numerous analytes, most of which were intended to determine binding to the carbohydrate antigen, in its unconjugated (Man-Gal), monomeric (Ep01-Man-Gal) or multivalent (FF03-Ep01-Man-Gal) forms, and others of which were intended to serve as test controls, such as the CRM197 protein. A scheme of the grid can be found in Figure 52.

	A	B	C	D	E	F
1	CRM 10 $\mu$ M	CRM 100 $\mu$ M	Ep01 10 $\mu$ M	Ep01 100 $\mu$ M	FF03-Ep01 10 $\mu$ M	FF03-Ep01 100 $\mu$ M
2		CRM 100 $\mu$ M	Ep01 10 $\mu$ M	Ep01 100 $\mu$ M	FF03-Ep01 10 $\mu$ M	FF03-Ep01 100 $\mu$ M
3	Man 10 $\mu$ M	Man 100 $\mu$ M	Ep01-Man 10 $\mu$ M	Ep01-Man 100 $\mu$ M	FF03-Ep01-Man 10 $\mu$ M	FF03-Ep01-Man 100 $\mu$ M
4	Man 10 $\mu$ M	Man 100 $\mu$ M	Ep01-Man 10 $\mu$ M	Ep01-Man 100 $\mu$ M	FF03-Ep01-Man 10 $\mu$ M	FF03-Ep01-Man 100 $\mu$ M
5	Man-Gal 10 $\mu$ M	Man-Gal 100 $\mu$ M	Ep01-Man-Gal 10 $\mu$ M	Ep01-Man-Gal 100 $\mu$ M	FF03-Ep01-Man-Gal 10 $\mu$ M	FF03-Ep01-Man-Gal 100 $\mu$ M
6	Man-Gal 10 $\mu$ M	Man-Gal 100 $\mu$ M	Ep01-Man-Gal 10 $\mu$ M	Ep01-Man-Gal 100 $\mu$ M	FF03-Ep01-Man-Gal 10 $\mu$ M	FF03-Ep01-Man-Gal 100 $\mu$ M

**Figure 52.** Representation of a microdot assay grid within a glass slide, indicating position and concentration of peptide/protein/carbohydrate.



On each grid we compared the *in vitro* immunogenicity of Man-Gal, Ep01-Man-Gal and FF03-Ep01-Man-Gal with the antibodies contained in the sera derived from five human patients and five canine subjects infected by the *Leishmania donovani* parasite. We employed four different antibody dilutions and two peptide/glycan concentrations. The concentrations were carefully selected to represent subsaturation and oversaturation levels to react with the N-hydroxyl succinimide (NHS) ester-activated glass slides. In this manner, the error that could derive from differences in the degree of immobilization of the peptides and carbohydrates was minimized. Moreover, we verified the validity of the microdot assay and the success of immobilization by including in our investigation the binding of ConA-FITC to the mannose presented by Ep01 and FF03-Ep01. Binding of anti-leishmanial antibodies from canine or human sources to the antigen was detected using fluorescently labelled secondary antibodies. The results of the microdot assay are summarized in Figure 53.



**Figure 53.** Anti- $\beta$ -Gal-(1-4)- $\alpha$ -Man antibody binding to unconjugated (Man-Gal), monomeric (Ep01-Man-Gal) and multivalent antigens (FF03-Ep01-Man-Gal), at the over-saturating concentration of 100  $\mu$ M. A: Mean fluorescence intensity of DyLight 488 indicating binding of canine anti- $\beta$ -Gal-(1 $\rightarrow$ 4)- $\alpha$ -Man IgG; B: Mean fluorescence intensity of AlexaFluor 647 indicating binding of human anti- $\beta$ -Gal-(1 $\rightarrow$ 4)- $\alpha$ -Man IgG; C: Mean fluorescence intensity of FITC indicating binding of ConA to 5-amino-pentanyl- $\alpha$ -D-mannopyranoside. Asterisks indicate P values (\*\*\*: P < 0.005; \*: P < 0.05). Under each column, a representative example of wells corresponding to the specific array is given (duplicates). 0: buffer only; 1: Man-Gal; 2: Ep01-Man-Gal; 3: FF03-Ep01-Man-Gal; 4: Man; 5: Ep01-Man; 6: FF03-Ep01-Man.

The microdot data with respect to ConA confirms the immobilization efficiency and the assay success: we observed significantly higher mannose recognition by ConA when the ligand is presented on the multivalent scaffold compared to its monovalent forms.

When the *Leishmania*-derived disaccharide is coupled to the Ep01 monomer, antibody recognition is comparable to that of the unconjugated antigen for both humans and dogs. This means that the conjugation of the antigen on the monovalent peptide Ep01 does not improve the recognition of the free antigen. Moreover, the epitope Ep01, which is by itself antigenic, does not alter the specificity of the anti-*Leishmania* antibodies.

The presentation of the *Leishmania* antigen Man-Gal on the multivalent peptide scaffold FF03 significantly enhances antibody binding by two-fold, in the case of canine sera, and by three-fold, for human sera compared to the monovalent display regime. The increased antibody binding observed with multivalent presentation of the carbohydrate antigen might be due to its high local concentration or/and to the regularity of its distribution on the scaffold.

These results demonstrate the potential of this coiled-coil-based self-assembling peptide as a scaffold for the presentation of peptide and carbohydrate antigens. The high local density of the presented antigens on the multivalent platform could be used to improve the sensitivity of numerous bioassays.

## 5.6 SUMMARY AND OUTLOOK II

The work described here was published within the framework of this thesis as an original research article entitled “*A Self-Assembling Peptide Scaffold for the Multivalent Presentation of Antigens*”.<sup>232</sup> The goal of this work was to characterize a peptide scaffold that would mimic the multivalent organization of native antigens and thereby enhancing antibody recognition of presented ligands. We have shown that a novel coiled coil-based, fiber-forming peptide can indeed act as a multivalent scaffold for the presentation of a peptide epitope and a carbohydrate antigen to their respective binding partners.

We have designed a fiber-forming peptide to increase local density of a highly immunogenic T-cell epitope (Ep01) from the diphtheria toxin (DT) and a carbohydrate antigen derived from the repeated lipophosphoglycan motif of the *Leishmania* parasite (Man-Gal). The presence of both T-cell and B-cell activators is the basic principle applied during the formulation of carbohydrate-based vaccines. As proof of concept, we generated a glycoconjugate presenting Ep01 and the monosaccharide mannose (Man). The self-assembly of the peptide scaffold here

called FF03 into long, non-branching fibers is achieved by placing hydrophobic amino acids at the sequence termini, to achieve the “sticky end” effect, and by alternating positive and negative charges on *b* and *c* positions of the coiled-coil heptad repeat, to induce interfibril interactions. Both Ep01 and Man/Man-Gal ligands were conjugated on the FF03 scaffold directly on solid phase. Once cleaved from the resin, the peptide self-assembled and the ligands were displayed at the most solvent-exposed positions. We verified the formation of the fibers by means of transmission electron microscopy (TEM), and we showed that FF03 retained its structural properties when loaded with peptide and carbohydrate ligands prior to assembly.

The efficient display of the ligands Ep01, Man and Man-Gal by the peptide FF03 was verified by means of ELISA, using antibodies produced by mice immunized against DT. The anti-DT antibodies recognized the specific peptide epitope more effectively when presented as a ligand on the self-assembled scaffold than in its unconjugated form. Moreover, at the micromolar range, FF03 presented the epitope Ep01 to anti-DT antibodies even more efficiently than the toxoid CRM197.

FF03 was investigated for its ability to present the carbohydrate ligand mannose to carbohydrate-binding proteins, both in solution, by employing the lectin concanavalin A, and as present on a bacterial cell surface, by means of a mannose-binding *E. coli* strain. In both cases, the mannose was optimally accessible to the lectins, which recognized it and bind it with high affinity.

To evaluate the ability of FF03 to efficiently present the antigenic carbohydrate Man-Gal on its surface, polyclonal antibodies against the *Leishmania* parasite from canine and human sources were employed in a microdot assay. Also in this case, the leishmanial carbohydrate antigen presented as ligand on the fiber-forming peptide scaffold was better recognized by specific antibodies in its multivalent form, rather than unconjugated.

Recent advances in immunology have demonstrated that simulating native antigen organization and recognition is an important design principle for nanomolecular structures for biomedical applications. Some potential applications of this multivalent scaffold are given below:

### **1. Fully synthetic vaccines**

The system presented in this work contains all the components of a synthetic carbohydrate-based vaccine. Fully synthetic vaccines are generally considered safer than vaccines produced by means of bacterial cell culture, they can be produced faster and at reduced costs.



Moreover, synthetic approaches allow for the production of a well-defined product that can be used not only for medicinal purposes but as tools to provide a better understanding of the molecular basis of immunity.

The here-described FF03-Ep01-Man-Gal could be further optimized to be tested for its ability to stimulate an immune response and generate immunological memory *in vivo*. The formation of the fibers should ensure higher stability to proteolytic degradation in the host organism; however, cytotoxicity experiment must be first carried out *in vitro*.

## 2. *In vitro* diagnostics

Typically, the faster a disease is correctly diagnosed, the more successfully it can be treated. In fact, certain medical conditions can only be considered curable if the diagnosis is formulated at an early stage. In a pattern recognition method, the key to diagnostic success is the identification of markers associated with the disease. Unfortunately, the outcome is often compromised by the poor availability of the marker; thus, highly sensitive assays are needed. The FF03 scaffold offers an attractive platform for presenting antigens at high densities. As demonstrated by the ELISA and microdot assays carried out over the course of this doctoral thesis, the peptide scaffold FF03 increases antibody avidity towards a T-cell epitope from the diphtheria toxin and a B-cell determinant from the *Leishmania* parasite. Its employment in diagnostics could increase the sensitivity of the assays.

## 3. Controlled release

The delivery of biologically-relevant compounds either upon response to stimuli or in a controlled manner over time generates numerous advantages. The controlled release of oral dose formulations, for example, prolongs the action while keeping the drug level within a safe therapeutic range, reducing unwanted side effects.

Due to the robustness of the scaffold presented in this thesis, it could be employed for the controlled release of therapeutics: for example, a drug could be conjugated to the scaffold via a short proteolytic amino acid sequence or via a pH-sensitive group. In the former case, the drug would be gradually released according to the availability of the scissile bond to the proteases; in the latter case, the drug would be released once it was incorporated into the lysosomes (pH 4.0) of the host cell. By adding to the scaffold a cell type-specific ligand (such as a binding partner for a receptor solely expressed on cancer cells), it would be possible to target a specific cellular subpopulation.

## 5.7 EXPERIMENTAL PROCEDURES II

### Peptide synthesis

Resins for solid-phase peptide synthesis were purchased from Novabiochem®. Preloaded Fmoc-Leu-NovaSyn® TGA resin and preloaded Fmoc-Gly-NovaSyn® TGT resin, both with 0.2 mmol/g amino acid loading, were used for the synthesis of FF03 and Ep01, respectively. The synthesis was performed with standard Fmoc/tBu chemistry, on a 0.05 mmol scale, using fluorenylmethyloxycarbonyl (Fmoc)-protected amino acids purchased from Orpegen. Removal of the Fmoc group at each step was achieved with 20% piperidine in DMF (2 x 10 minutes and 2 x 5 minutes with 5 mL deprotection solution).

The synthesis of Ep01 was performed manually with single 1-hour couplings, using an eight-fold excess of amino acid, HOAt and DIC, relative to resin loading. Manual synthesis was carried out to avoid histidine racemization.

The synthesis of FF03 and FF03-Ep01 was performed in two main steps: FF03 was synthesized with 2-hour double couplings using a SyroXP-I peptide synthesizer (Multi-SynTech GmbH), with an eight-fold excess of amino acid, TBTU/HOBt as coupling agents and DIPEA, relative to resin loading. The lysine on position 17 was protected with N-methyltrityl (Mtt), an amine protective group removable under mild acidic conditions. After synthesizing the full-length FF03 sequence and blocking the N-terminus with Boc-Abz-OH, selective deprotection of K17 was performed (1% TFA in DCM). This free amine was used as a starting point for the orthogonal synthesis of Ep01. The synthesis of mannose derivative 5-amino-pentanyl- $\alpha$ -D-mannopyranoside and of the 5-amino-pentanyl  $\beta$ -D-galactopyranosyl-(1 $\rightarrow$ 2)- $\alpha$ -D-mannopyranoside, derived from the leishmanial lipophosphoglycan, was performed by our collaborators at the MPI as previously reported. The addition of the carbohydrate moieties to Ep01 and FF03-Ep01 was carried out by first switching the functionality on resin from  $-\text{NH}_2$  to  $-\text{COOH}$  with a five-fold excess of glutaric anhydride and a catalytic amount of DIPEA for 3 hours. The further activation of the free acid with 3 equivalents of HATU and 6 equivalents of DIPEA for 5 minutes allowed coupling of the sugars “on resin”, using a three-fold excess of building block. Final cleavage from the resin was performed using a solution of 95% TFA, 3% water and 2% TIS (5 mL) with shaking for 3 hours. The resin was washed with 1 mL TFA and 1 mL DCM and excess solvent was removed by evaporation. The peptides were precipitated in cold diethyl ether and collected by centrifugation.

## **Peptide purification**

The peptides were purified via analytical HPLC. This system was preferred over preparative HPLC because of the difficulty inherent in separating the main products from the side products (monoglycosylation did not lead to a significant difference in elution time). Thus, purification was performed on a LaChrom-ELITE-HPLC-System (VWR) consisting of two HPLC-pumps (L-2130) with a solvent degasser, an auto sampler (L-2200), a diode flow detector (L-2455) and a high-pressure gradient mixer. Data analysis was performed with EZ Chrom ELITE software. All runs were performed with a flow rate of 1.0 mL/min using acetonitrile (ACN, 0.1% TFA) and Millipore H<sub>2</sub>O (0.1% TFA).

## **Peptide concentration determination**

Peptide concentration was determined by UV spectroscopy using the absorption maximum at 280 nm in 6M guanidine hydrochloride, as previously reported.<sup>233</sup> Peptide FF03 concentration was determined by a standard curve of Abz absorbance at 320 nm.

## **Peptide characterization**

Pure peptides were characterized via mass spectroscopy (MS), circular dichroism (CD) and transmission electron microscopy (TEM).

The quality of the synthesized glycopeptides was determined using an ESI-TOF 6210 from Agilent (USA, CA-95051-7201, Santa Clara). All samples were dissolved in a mixture of water and ACN before injection and the data were collected on positive ion mode.

For the CD experiments, peptides were dissolved in 100 mM phosphate buffer and the pH was adjusted to 7.4 with either NaOH or HCl. Spectra were recorded on a JASCO-8-10 spectropolarimeter with a temperature controller set at 20 °C. The CD data were registered using 0.1 cm path length quartz cuvettes and normalized according to extinction coefficient, optical path length, peptide concentration and number of residues. During all measurements, a constant N<sub>2</sub> flush of 3.0 L/min was provided.

Samples for transmission electron microscopy were prepared by absorbing 5 μL aliquots of peptide solution onto glow-discharged carbon-coated collodium films on 400-mesh copper grids. The grids were blotted, stained with 1% phosphotungstic acid (PTA), and air dried.

TEM micrographs were taken at a primary magnification of 58300 using a defocus of 0.8 μm.

## **ELISA**

1  $\mu$ M stock solutions of Ep01 and FF03-Ep01 were prepared in PBS and incubated overnight to allow completion of self-assembly. Samples (50  $\mu$ L) were immobilized on 96-well Nunclon flat-bottom transparent polystyrene plates at three different concentrations (1  $\mu$ M, 100 nM, 10 nM) in triplicate. As a positive control, the Diphtheria protein CRM197 (MBL, RK-01-515) was immobilized in triplicate at the same concentrations. After incubation, the supernatant was aspirated and 50  $\mu$ L of blocking solution BPBS (1% BSA in PBS) was added to the wells and incubated for 1 hour at 37° C. The wells were then washed 3 times with PBS containing 0.05% tween). A mix of mouse polyclonal sera against CRM197 was added to the wells at five different dilutions, in ten-fold steps from 1:1000 to 1:10000000, and left to incubate for 1 hour and 30 minutes at 37° C. Details about immunization can be found in the publication of our collaborators.<sup>200</sup> After aspiration of supernatants, the wells were washed for 15 minutes with 150 mM NH<sub>4</sub>SCN, then PBST (x2). Goat anti-Mouse IgG Fc antibody, Horseradish Peroxidase (HRP)-conjugate (Dianova), diluted 1:20000 in PBS (50  $\mu$ L), was allowed to incubate for 1 hour at 37° C and then removed. After washing again with PBST (x3), 50  $\mu$ L of 3,3',5,5' tetramethylbenzidine peroxidase substrate solution (1-Step Ultra TMB-ELISA) was added and left to react for 3 minutes. The reaction was stopped by addition of 2% sulfuric acid (50  $\mu$ L). The plates were then read 3 times at 450 nm with the multimode microplate reader Infinite® 200 PRO by Tecan. Analysis was performed using Microsoft Excel software for data plotting and *p* value calculation (Student T-test type and the number of tails: 1).

## **Confocal laser microscopy (CLM)**

Fluorescent images were acquired with the instrument Zeiss LSM 700, using the laser at a wavelength of 488 nm (2.0 %) for FITC excitation. The selected objective was a Plan Achromat 63X/1.4 oil DIC and the images were processed with the software ZEN 2009 by Zeiss.

## **ConA-FITC binding assay**

The two mannose-containing peptides and their controls without sugar (1.5 mg/mL) were dissolved in lectin-binding buffer (10 mM HEPES, 1 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, pH 7.4) and

spotted twice on poly-lysine-coated microscope slides. The samples were left overnight to self-assemble and 15  $\mu$ L of blocking solution (1% BSA in PBS) was added to each spot and incubated for 1 hour. One drop of ConA from *Canavalia ensiformis*, FITC-conjugated (5  $\mu$ L, 5 mg/mL, purchased from Sigma-Aldrich) in lectin-binding buffer was added and left for 1 hour in the dark to interact with the peptides. The spots were washed twice with 15  $\mu$ L PBS, and one drop of buffer was added before observing the results with the CLM.

### **Mannose-binding *E. coli* assay**

*E. coli* strain ORN178 was grown in Luria-Bertani broth in the presence of 12.4  $\mu$ g/mL tetracycline at 37 °C with shaking to an OD<sub>595nm</sub> of 0.8. 50  $\mu$ L culture was added to a 96-well plate (in triplicate) and the cells were collected by centrifugation (4°C, 500 X g, 20 minutes). The bacteria were washed three times with ice-cold PBS and incubated for 1 h at room temperature with 50  $\mu$ L PBS containing FITC (500 nmol final concentration). The cells were again washed three times and 20  $\mu$ L of the suspension was spotted onto polylysine microscope plates upon which 20  $\mu$ L of a 1.5 mg/mL solution of FF03-Ep01 or FF03-Ep01-Man in PBS had been left overnight to self-assemble. After incubation at room temperature, the cells were observed with CLSM.

### **Microdot assay**

The peptide library, CRM197 and the carbohydrates were immobilized on CodeLink N-hydroxyl succinimide (NHS) ester activated glass slides (SurModics Inc., Eden Prairie) with a piezoelectric spotting device (S3; Scienion). The subsaturation concentration of 10  $\mu$ M and the oversaturation concentration of 100  $\mu$ M of compounds to react with NHS groups on the slide were chosen to minimize the effect of difference in antibody reactivity due to possible disparity in immobilization levels. On each slide, 64 replicate array grids were printed. The slides were incubated in a humid chamber for 24 h until reaction completion and quenched with 100 mM ethanolamine in 50 mM sodium phosphate buffer at pH 9 for 1 h at 50 °C. A washing procedure (x3) with deionized water followed and the slides were dried via centrifugation for 5 minutes at 1200 X g. Before use, the slides were blocked with 1% BSA in PBS overnight at 4 °C, washed 3 times with PBS and dried again. With the help of the 64-well grid (ProPlate multi-array system), two different slides were treated to test the reactivity

of the primary antibodies from sera obtained from five human patients and five canine subjects infected with the Leishmania parasite. Details about immunization procedures can be found in the publication of our collaborators from MPI. The slides were incubated with 40  $\mu$ L of four dilutions (1:100, 1:200, 1:400 and 1:800) of the appropriate serum in blocking buffer (20  $\mu$ L) for 1 h at 37 °C. The slides were washed 3 times with 40  $\mu$ L PBST and incubated for 1 h at 37 °C with 40  $\mu$ L of secondary-antibody solution in BPBS. The secondary antibodies used are as follows: AlexaFluor 647 goat  $\alpha$ -human IgG (Life Technologies) for the human sera, AlexaFluor 635 goat  $\alpha$ -mouse IgG (Invitrogen) for the mouse sera (in the case of the CRM197 control), DyLight 488 rabbit  $\alpha$ -dog IgG (Fuller Laboratories) for the canine sera. The slides were washed 4 times with 40  $\mu$ L PBST and dried as previously before scanning with a GenePix 4300A scanner (Bucher Biotec). Excitation wavelengths of 488 nm and 633 nm were used for FITC and DyLight 488, and AlexaFluor 647 and AlexaFluor 635, respectively; emission was detected by means of filters set to either 530 nm or 660 nm, respectively. The data were acquired with GenePix Pro 7 software (Bucher Biotec). Data and *p* values calculation were performed using Microsoft Excel software. Student T-test type and number of tails selected were 1 and 2, respectively.

# 6

**Synthetic Glycopeptides as Multivalent Scaffolds for Carbohydrates:**  
from receptor targeting to vaccines exploiting sugar-protein interactions

## REFERENCES

## REFERENCES

---

- <sup>1</sup> Arcuate Glucagon-Like Peptide 1 Receptors Regulate Glucose Homeostasis but Not Food Intake. Sandoval D., Bagnol D., Woods S., D'Alessio D. and Seeley R. *Diabetes* **2008**, *57*, 2056-2054.
- <sup>2</sup> Homeostatic Regulation of T Cell Trafficking by a B Cell-Derived Peptide is Impaired in Autoimmune and Chronic Inflammatory Disease. Chimen M., McGettrick H., Apta B., Kuravi S., Yates C., Kennedy A., Odedra A., Alassiri M., Harrison M., Martin A., Narendran P., Rainger G. *et al. Nature Med* **2015**, *21*, 467-475.
- <sup>3</sup> Coupled folding and binding with alpha-helix-forming molecular recognition elements. Oldfield C., Cheng Y., Cortese M., Romero P., Uversky V., Dunker A. *Biochem* **2005**, *44*(37), 12454-70.
- <sup>4</sup> Importance of the fourth alpha-helix within the CAP homology domain of type II topoisomerase for DNA cleavage site recognition and quinolone action. Strumberg D., Nitiss J., Dong J., Walker J., Nicklaus M., Kohn K., Heddle J., Maxwell A., Seeber S., Pommier Y. *Antimicrob Agents Chemother.* **2002**, *46*(9), 2735-46.
- <sup>5</sup> Proposed Mechanism for the Biological Assembly of Collagen Triple Helix. Speakman, P. T. *Nature* **1971**, *229*(5282), 241-243.
- <sup>6</sup> The Location of the Repressor Binding Sites in the lac Operon. Reznikoff W., Winter R. and Katovich C. *Hurley Proc Natl Acad Sci U S A* **1974**, *71*(6), 2314-2318.
- <sup>7</sup> Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis. Sattler M., Liang H., Nettesheim D., Meadows R., Harlan J., Eberstadt M., Yoon H., Shuker S., Chang B., Minn A., Thompson C., Fesik S. *Science* **1997**, *275*(5302), 983-6.
- <sup>8</sup> Electric field effects on membranes: gramicidin A as a test ground. Siu S., Böckmann R. *J Struct Biol* **2007**, *157*(3), 545-556.
- <sup>9</sup> Carbohydrate and glycoprotein specificity of two endogenous cerebellar lectins. Marschal P., Reeber A., Neeser J., Vincendon G., Zanetta J. *Biochimie* **1989**, *71*(5), 645-53.
- <sup>10</sup> Molecular mechanisms of HIV entry. Wilen C., Tilton J., Doms R. *Viral Molecular Machines, Advances in Experimental Medicine and Biology* **2012**, *726*, 223-242.



- 
- <sup>11</sup> The Future of Peptide-based Drugs. Craik D., Fairlie D., Liras S. and Price D. *Chem Biol Drug Des* **2013**, *81*, 136-147.
- <sup>12</sup> Transparent Market Research, press release. Peptide Therapeutics Market (by Applications, by Route of Administration, and by Marketing Status) - Global Industry Analysis, Size, Share, Growth, Trends and Forecast 2014 – 2020.
- <sup>13</sup> Pharmaceutical applications of bioactive peptides. Danquah M. and Agyei D. *Biotechnology* **2012**, *1(2)*, 5.
- <sup>14</sup> Radiolabelled antimicrobial peptides for infection detection. Lupetti A., Welling M., Pauwels E., Nibbering P. *Lancet Infect Dis.* **2003**, *3(4)*, 223-229.
- <sup>15</sup> Synthetic peptides in the diagnosis of neurological diseases. Lolli F., Mazzanti B., Rovero P., Papini A. *Curr Protein Pept Sci.* **2003**, *4(4)*, 277-284.
- <sup>16</sup> Metabolic actions of natriuretic peptides and therapeutic potential in the metabolic syndrome. Schluetera N., de Sterkea A., Willmesa D., Spranger J., Jordanb J., Birkenfeld A. *Pharmacol Ther* **2014**, *144(1)*, 12–27.
- <sup>17</sup> Mechanisms of Protease Action. Polgar L. *Proteases and Regulation*, CRC Press **1989**, Chapter 3.
- <sup>18</sup> Biodegradable polymeric nanoparticles based drug delivery systems. Kumari A., Kumar Yadav S., Yadav S. *Colloids Surf B* **2010**, *75(1)*, 1-18.
- <sup>19</sup> Cationic liposomes as efficient nanocarriers for the drug delivery of an anticancer cholesterol-based ruthenium complex. Vitiello G., Luchini A., D'Errico G., Santamaria R., Capuozzo A., Irace C., Montesarchio D. and Paduano L. *J Mater Chem B* **2015**, *3*, 3011-3023.
- <sup>20</sup> Polymeric Micelles, a Promising Drug Delivery System to Enhance Bioavailability of Poorly Water-Soluble Drugs. Xu W., Ling P. and Zhang T. *Drug Deliv* **2013**, *2013*, 340315.
- <sup>21</sup> Stealth Properties to Improve Therapeutic Efficacy of Drug Nanocarriers. Salmaso S. and Caliceti P. *Drug Deliv* **2013**, *2013*, 374252.
- <sup>22</sup> Recent development of peptide self-assembly. Zhao X., Pan F., Lu J. *Prog Nat Sci* **2008**, *18(6)*, 653–660.

- 
- <sup>23</sup> Chemomechanical mapping of ligand–receptor binding kinetics on cells. Lee S., Mandic J. and Van Vliet K. *PNAS* **2007**, *104*(23), 9609–9614.
- <sup>24</sup> Multivalency as a Chemical Organization and Action Principle. Fasting C., Schalley C., Weber M., Seitz O., Hecht S., Kokschi B., Dornedde J., Graf C., Knapp E. and Haag R. *Angew Chem Int Ed* **2012**, *51*, 10472-10498.
- <sup>25</sup> Synthetic Biomaterials as Instructive Extracellular Microenvironments for Morphogenesis in Tissue Engineering. Lutolf M., Hubbell J. *Nat Biotechnol* **2005**, *1*, 47-55.
- <sup>26</sup> Selective Differentiation of Neural Progenitor Cells by High-Epitope Density Nanofibers. Silva G., Czeisler C., Niece K., Beniash E., Harrington D., Kessler J., Stupp S. *Science* **2004**, *303*, 1352-1355.
- <sup>27</sup> Self-Assembling Peptide Nanofiber Scaffolds Accelerate Wound Healing. Schneider A., Garlick J., Egles C. *PLoS ONE* **2008**, *3*, e1410.
- <sup>28</sup> Engineering Pro-Angiogenic Peptides Using Stable, Disulfide-Rich Cyclic Scaffolds. Chan L., Gunasekera S., Henriques S., Worth N., Le S., Clark R., Campbell J., Craik D. and Daly N. *Blood* **2011**, *118*, 6709-17.
- <sup>29</sup> Multivalent glycoconjugates As Anti-Pathogenic Agents. Bernardi A., Jiménez-Barbero J., Casnati A., De Castro C., Darbre T., Fieschi F., Finne J., Funken H., Jaeger K., Lahmann M. *et al. Chem. Soc. Rev.* **2013**, *42*, 4709-4727.
- <sup>30</sup> Quantitative Investigation of the Modular Primer Effect for DNA and Peptide Nucleic Acid Hexamers. Nilsson P., O'meara D., Edebratt F., Persson B., Uhlén M., Lundeberg J., Nygren P. *Anal Biochem.* **1999**, *269*, 155-161.
- <sup>31</sup> Peptide-LNA Oligonucleotide Conjugates. Astakhova I., Hansen L., Vester B., Wengel J. *Org. Biomol. Chem.* **2013**, *11*, 4240-4249.
- <sup>32</sup> Sequence dependence of kinetics and morphology of collagen model peptide self-assembly into higher order structures. Karunakar K., Wang Y. and Brodsky B. *Protein Sci* **2008**, *17*(6), 1086-1095.
- <sup>33</sup> Inducible proteopathies. Walker L., LeVine H., Mattson M. and Jucker M. *Trends Neurosci* **2006**, *29*(8), 438-443.

- 
- <sup>34</sup> From natural to designed self-assembling biopolymers, the structural characterisation of fibrous proteins & peptides using fibre diffraction. Morriss K. and Serpell L. *Chem Soc Rev* **2010**, *39*, 3445–3453.
- <sup>35</sup> Thermal and Chemical Stability of Diphenylalanine Peptide Nanotubes: Implications for Nanotechnological Applications. Adler-Abramovich L., Reches M., Sedman V., Allen S., Tendler S. and Gazit E. *Langmuir* **2006**, *22*(3), 1313-1320.
- <sup>36</sup> New Cyclic Peptide Assemblies with Hydrophobic Cavities: The Structural and Thermodynamic Basis of a New Class of Peptide Nanotubes. Amorín M., Castedo L. and Granja J. *J Am Chem Soc* **2003**, *125*(10), 2844-2845.
- <sup>37</sup> Molecular and cellular aspects of protein misfolding and disease. Herczenik E. and Gebbink M. *FASEB J* **2008**, *22*(7), 2115-2133.
- <sup>38</sup> Functional amyloid formation within mammalian tissue. Fowler D., Koulov A., Alory-Jost C., Marks M., Balch W. and Kelly J. *PLoS Biol.* **2006**, *4*(1), e6.
- <sup>39</sup> The coiled-coil model of  $\alpha$ -keratin structure. Fraser R., MacRae T., Miller A. *J Mol Biol* **1964**, *10*(1), 147-156.
- <sup>40</sup> Myosin structure and function in cell motility. Warrick H. and Spudich J. *Ann Rev Cell Biol* **1987**, *3*, 379-421.
- <sup>41</sup> Protein transduction: cell penetrating peptides and their therapeutic applications. Wagstaff K. and Jans D. *Curr Med Chem.* **2006**, *13*(12), 1371-87.
- <sup>42</sup> Mechanisms of Cellular Uptake of Cell-Penetrating Peptides. Madani F., Lindberg S., Langel Ü., Futaki S. and Gräslund A. *J Biophys* **2011**, *2011*:414729.
- <sup>43</sup> A peptide carrier for the delivery of biologically active proteins into mammalian cells. Morris M., Depollier J., Mery J., Heitz F. and Divita G. *Nat Biotechnol* **2001**, *19*, 1173–1176.
- <sup>44</sup> Peptide-based delivery of nucleic acids: design, mechanism of uptake and applications to splice-correcting oligonucleotides. Abes S., Moulton H., Clair P., Prevot P., Youngblood D., Wu R., Iversen P. and Lebleu B. *J Controlled Release*, **2007**, *116*, 304-313.

- 
- <sup>45</sup> Efficient splicing correction by PNA conjugation to an R6-Penetratin delivery peptide. Abes S., Turner J., Ivanova G., Owen D., Williams D., Arzumanov A., Clair P., Gait M. and Lebleu B. *Nucleic Acids Res* **2007**, *35*, 4495-4502.
- <sup>46</sup> BMAP-28, an Antibiotic Peptide of Innate Immunity, Induces Cell Death through Opening of the Mitochondrial Permeability Transition Pore. Risso A., Braidot E., Sordano M., Vianello A., Macri F., Skerlavaj B., Zanetti M., Gennaro R. and Bernardi P. *Mol Cell Biol* **2002**, *22*, 1926-1935.
- <sup>47</sup> TAT conjugated, FITC doped silica nanoparticles for bioimaging applications. Santra S., Yang H., Dutta D., Stanley J., Holloway P., Tan W., Moudgil B. and Mericle R. *Chem Commun* **2004**, *24*, 2810–2111.
- <sup>48</sup> A synthetic peptide mediated active targeting of cisplatin liposomes to Tie2 expressing cells. Mai J., Song S., Rui M., Liu D., Ding Q., Peng J., Xu Y. *J Controlled Release* **2009**, *139*, 174-181.
- <sup>49</sup> Peptide decoration of nanovehicles to achieve active targeting and pathology-responsive cellular uptake for bone metastasis chemotherapy. Wang X., Yang Y., Jia H., Jia W., Miller S., Bowman B., Feng J. and Zhan F. *Biomater Sci*, **2014**, *2*, 961.
- <sup>50</sup> Peptide-based targeting strategies for simultaneous imaging and therapy with nanovectors. Accardo A., Tesauro D. and Morelli G. *Polymer Journal* **2013**, *45*, 481-493.
- <sup>51</sup> Nuclear targeting peptide scaffolds for lipofection of nondividing mammalian cells. Subramanian A., Ranganathan P. and Diamond S. *Nat Biotech* **1999**, *17*, 873-877.
- <sup>52</sup> Controlling hydrogelation kinetics by peptide design for three-dimensional encapsulation and injectable delivery of cells. Haines-Butterick L., Rajagopal K., Branco M., Salick D., Rughani R., Pilarz M., Lamm M., Pochan D. and Schneider J. *PNAS* **2014**, *104*(19), 7791-7796.
- <sup>53</sup> Self-assembled peptide-based hydrogels as scaffolds for anchorage-dependent cells. Zhou M., Smith A., Das A., Hodson N., Collins R., Ulijn R., Gough J. *Biomaterials* **2009**, *30*(13), 2523-2530.

- 
- <sup>54</sup> Self-Assembly and Mineralization of Peptide-Amphiphile Nanofibers. Hartgerink J., Beniash E., Stupp S. *Science* **2001**, 294(5547), 1684-1688.
- <sup>55</sup> Designed  $\alpha$ -Helical Tectons for Constructing Multicomponent Synthetic Biological Systems. Bromley E., Sessions R., Thomson A. and Woolfson D. *J Am Chem Soc* **2009**, 131(3), 928-930.
- <sup>56</sup> Supramolecular design of self-assembling nanofibers for cartilage regeneration. Shah R., Shah N., Del Rosario Lim M., Hsieh C., Nuber G., Stupp S. *Proc Natl Acad Sci USA* **2010**, 107(8), 3293-3298.
- <sup>57</sup> BMHP1-derived self-assembling peptides: hierarchically assembled structures with self-healing propensity and potential for tissue engineering applications. Gelain F., Silva D., Caprini A., Taraballi F., Natalello A., Villa O., Nam K., Zuckermann R., Doglia S., Vescovi A. *ACS Nano* **2011**, 5(3), 1845-1859.
- <sup>58</sup> Self-Assembling Peptides as Cell-Interactive Scaffolds. Wu E., Zhang S. and Hauser C. *Adv Funct Mater* **2011**, 22(3), 456-468.
- <sup>59</sup> A computationally directed screen identifying interacting coiled coils from *Saccharomyces cerevisiae*. Newman J., Wolf E. and Kim P. *Proc Natl Acad Sci* **2000**, 97, 13203-13208.
- <sup>60</sup> Thermodynamic characterization of the structural stability of the coiled-coil region of the bZIP transcription factor GCN4. Thompson K., Vinson C., Freire E. *Biochemistry* **1993**, 32(21), 5491-5496.
- <sup>61</sup> The myosin coiled-coil is a truly elastic protein structure. Schwaiger I., Sattler C., Hostetter D. and Rief M. *Nat Materials* **2002**, 1, 232-235.
- <sup>62</sup> Structural insights into the SNARE mechanism. Fasshauer D. *Bioch Bioph Acta* **2003**, 1641, 87-97.
- <sup>63</sup> GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. Hope I. and Struhl K. *EMBO J.* **1987**, 6(9), 2781-2784.
- <sup>64</sup> Fos and Jun: oncogenic transcription factors. Curran T. *Tohoku J Exp Med* **1992**, 168(2), 169-74.

- 
- <sup>65</sup> Structural and Functional Insights on the Myosin Superfamily. Syamaladevi D., Spudich J. and Sowdhamini R. *Bioinform Biol Insights* **2012**, *6*, 11-21.
- <sup>66</sup> Is  $\alpha$ -Keratin a Coiled Coil? Crick F. *Nature* **1952**, *170*, 882-883.
- <sup>67</sup> The packing of alpha-helices: simple coiled-coils. Crick F. *Acta Cryst* **1953**, *6*, 689-697.
- <sup>68</sup> A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. Harbury P., Zhang T., Kim P. and Alber T. *Science* **1993**, *262*(5138), 1401-1407.
- <sup>69</sup> The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Landschulz W., Johnson P., McKnight S. *Science* **1988**, *240*, 1759-1764.
- <sup>70</sup> Coiled coils: new structures and new functions. Lupas A. *Trends Biochem Sci* **1996**, *21*, 375-382.
- <sup>71</sup> A network of coiled-coil associations derived from synthetic GCN4 leucine-zipper arrays. Portwich M., Keller S., Strauss H., Mahrenholz C., Kretzschmar I., Kramer A., Volkmer R. *Angew Chem Int Ed Engl* **2007**, *46*(10),1654-1657.
- <sup>72</sup> The role of position *a* in determining the stability and oligomerization state of  $\alpha$ -helical coiled coils: 20 amino acid stability coefficients in the hydrophobic core of proteins. Wagschal K., Tripet B., Lavigne P., Mant C. and Hodges R. *Protein Science* **1999**, *8*, 2312-2329.
- <sup>73</sup> Leucine Zippers. Hakoshima, T. *eLS*, **2005**.
- <sup>74</sup> Predicting Helix Orientation for Coiled-Coil Dimers. Apgar J., Gutwin K. and Keating A. *Proteins* **2008**, *72*, 1048-1065.
- <sup>75</sup> How Post-Translational Modifications Influence amyloid Formation: A Systematic Study of Phosphorylation and Glycosylation in Model Peptides. Falenski J., Broncel M., Wagner S., Hackenberger C. and Kokschi B. *Chem. Eur. J.* **2010**, *16*, 7881-7888.
- <sup>76</sup> Nanoparticle-Induced Peptide Folding and Aggregation. Wagner S., Roskamp M., Pallerla M., Araghi R., Schlecht S. and Kokschi B. *Small* **2010**, *6*, 1321-1328.

- 
- <sup>77</sup> Towards Understanding Secondary Structure Transitions: Phosphorylation and Metal Coordination in Model Peptides. Broncel, M.; Wagner, S. C.; Hackenberger, C.; Kokscha, B. *Org. Biomol. Chem.* **2010**, *8*, 2575-2579.
- <sup>78</sup> Coiled Coil Domains: Stability, Specificity, and Biological Implications. Mason M. and Arndt K. *Chem Bio Chem* **2004**, *5*, 170-176.
- <sup>79</sup> X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. O'Shea E., Klemm J., Kim P. and Alber T. *Science* **1991**, *254*(5031), 539-544.
- <sup>80</sup> Native-like and structurally characterized designed alpha-helical bundles. Betz S., Bryson J. and DeGrado W. *Curr Opin Struct Biol.* **1995**, *5*(4), 457-463.
- <sup>81</sup> Leucine is the most stabilizing aliphatic amino acid in the d position of a dimeric leucine zipper coiled coil. Moitra J., Szilák L., Krylov D. and Vinson C. *Biochemistry* **1997**, *36*(41), 12567-12573.
- <sup>82</sup> Protein destabilization by electrostatic repulsions in the two-stranded alpha-helical coiled-coil/leucine zipper. Kohn W., Kay C. and Hodges R. *Protein Sci.* **1995**, *4*(2), 237-250.
- <sup>83</sup> Coiled coils: attractive protein folding motifs for the fabrication of self-assembled, responsive and bioactive materials. Apostolovic B., Danial M. and Klok HA. *Chem. Soc. Rev.* **2010**, *39*, 3541-3575.
- <sup>84</sup> Assembly pathway of a designed alpha-helical protein fiber. Bromley E., Channon K., King P., Mahmoud Z., Banwell E., Butler M., Crump M., Dafforn T., Hicks M., Hirst J., Rodger A. and Woolfson D. *Bioph J* **2010**, *98*(8), 1668-1676.
- <sup>85</sup> Rational design and application of responsive alpha-helical peptide hydrogels. Banwell E., Abelardo E., Adams D., Birchall M., Corrigan A., Donald A., Kirkland M., Serpell L., Butler M., Woolfson D. *Nat Mater.* **2009**, *8*(7), 596-600.
- <sup>86</sup> Biotechnological Production of Spider-Silk Proteins Enables New Applications. Vendrely C. and Scheibel T. *Macromol Biosc.* **2007**, *7*(4), 401-409.
- <sup>87</sup> Self-assembled peptide nanostructures: the design of molecular building blocks and their technological utilization. Gazit E. *Chem Soc Rev.* **2007**, *36*, 1263-1269.

- 
- <sup>88</sup> Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. King N., Sheffler W., Sawaya M., Vollmar B., Sumida J., André I., Gonen T., Yeates T. and Baker D. *Science* **2012**, *336*, 1171-1174.
- <sup>89</sup> Sticky Egyptians: A technique for assembling genes encoding constrained peptides of variable length. Palmer S., Redfern M., Smith G. and Cox J. *Nucl Acids Res* **1998**, *26*(11), 2560-2564.
- <sup>90</sup> Sticky-End Assembly of a Designed Peptide Fiber Provides Insight into Protein Fibrillogenesis. Pandya M., Spooner G., Sunde M., Thorpe J., Rodger A. and Woolfson D. *Biochemistry* **2000**, *39*, 8728-8734.
- <sup>91</sup> Introducing Branches into a Self-Assembling Peptide Fiber. Ryadnov M. and Woolfson D. *Angew Chem Int Ed.* **2003**, *42*, 3021- 3023.
- <sup>92</sup> De novo design of fibrils made of short alpha-helical coiled coil peptides. Potekhin S., Melnik T., Popov V., Lanina N., Vazina A., Rigler P., Verdini A., Corradin G. and Kajava A. *Chem Biol.* **2001**, *8*(11),1025-32.
- <sup>93</sup> Engineering nanoscale order into a designed protein fiber. Papapostolou D., Smith A., Atkins E., Oliver S., Ryadnov M., Serpell L. and Woolfson D. *PNAS* **2007**, *104*(26), 10853-10858.
- <sup>94</sup> Formation of  $\alpha$ -Helical Nanofibers by Mixing  $\beta$ -Structured and  $\alpha$ -Helical Coiled Coil Peptides. Brandenburg E., v. Berlepsch H., Leiterer J., Emmerling F and Kokschi B. *Biomacromolecules* **2012**, *13*(11), 3542-3551.
- <sup>95</sup> The use of peptidic frameworks for the construction of molecular receptors and devices. Voyer N. and Lamothe J. *Tetrahedron* **1995**, *51*(34), 9241-9284.
- <sup>96</sup> Conformations and free energy landscapes of polyproline peptides. Moradi M., Babin V., Roland C., Darden T. and Sagui C. *PNAS* **2009**, *106*(49), 20746-20751.
- <sup>97</sup> Cell Penetrating Agents Based on a Polyproline Helix Scaffold. Fillon Y., Anderson J. and Chmielewski J. *J Am Chem Soc.* **2005**, *127*, 11798-11803.



- 
- <sup>98</sup> Chemical and enzymatic synthesis of multivalent sialoglycopeptides. Unverzagt C., Kelm S. and Paulson J. *Carbohydr Res.* **1994**, *251*, 285-301.
- <sup>99</sup> Distinguishing Helix Conformations in Alanine-Rich Peptides Using the Unnatural Amino Acid TOAC and Electron Spin Resonance. Hanson P., Martinez G., Millhauser G, Formaggio F., Crisma M., Toniolo C. and Vita C. *J Am Chem Soc.* **1996**, *118*(1), 271-272.
- <sup>100</sup> Synthesis and characterization of water-soluble silk peptides and recombinant silk protein containing polyalanine, the integrin binding site, and two glutamic acids at each terminal site as a possible candidate for use in bone repair materials. Asakura T., Suzuki Y., Nagano A., Knight D., Kamiya M. and Demura M. *Biomacromolecules* **2013**, *14*(10), 3731-41.
- <sup>101</sup> Architecture Effects on the Binding of Cholera Toxin by Helical Glycopolypeptides. Liu C. and Kiick K. *Macromolecules* **2008**, *41*(3), 764-772.
- <sup>102</sup> Multiple glycosylation of de novo designed  $\alpha$ -helical coiled coil peptides. Falenski J., Gerling U and Kokschi B. *Bioorg Med Chem.* **2010**, *18*, 3703-3706.
- <sup>103</sup> The Cluster Glycoside Effect. Lundquist J. and Toone E. *Chem Rev.* **2002**, *102*(2), 555-578.
- <sup>104</sup> Glycoproteins. Spiro R. *Adv Protein Chem.* **1973**, *27*, 349-467.
- <sup>105</sup> Glycoscience 2nd ed. Wittmann V., Fraser-Reid B., Tatsuta K., Thiem J. *Springer-Verlag* **2008**.
- <sup>106</sup> Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Spiro R. *Glycobiology* **2002**, *12*(4), 43R-56R.
- <sup>107</sup> Site-selective chemical protein glycosylation protects from autolysis and proteolytic degradation. Russell D., Oldham N. and Davis B. *Carb Res.* **2009**, *344*, 1508-1514.
- <sup>108</sup> Conformational influences of glycosylation of a peptide: A possible model for the effect of glycosylation on the rate of protein folding. Live D., Kumar R., Beebe X. and Danishefsky S. *PNAS* **1996**, *93*(23), 12759-12761.
- <sup>109</sup> Cycling of O-linked bold beta-N-acetylglucosamine on nucleocytoplasmic proteins. Hart G., Housley M. and Slawson C. *Nature* **2007**, *446*, 1017-1022.

- 
- <sup>110</sup> Glycomics Hits the Big Time. Hart G. and Copeland R. *Cell* **2010**, *143*(5), 672-676.
- <sup>111</sup> Human Disease Glycomics/Proteome Initiative (HGPI). Taniguchi N. *Mol Cell Proteomics* **2008**, *7*(3), 626-627.
- <sup>112</sup> Fringe is a glycosyltransferase that modifies Notch. Moloney D., Panin V., Johnston S., Chen J., Shao L., Wilson R., Wang Y., Stanley P., Irvine K., Haltiwanger R. and Vogt T. *Nature* **2000**, *406*(6794), 369-375.
- <sup>113</sup> Lattices, rafts, and scaffolds: domain regulation of receptor signaling at the plasma membrane. Lajoie P., Goetz J., Dennis J. and Nabi I. *J Cell Biol.* **2009**, *185*(3), 381-385.
- <sup>114</sup> Evidence for autoimmune antibodies directed against embryonic neural cell adhesion molecules (N-CAM) in patients with group B meningitis. Nedelec J., Boucraut J., Garnier J., Bernard D. and Rougon G. *Neuroimmunol.* **1990**, *29*, 49-56.
- <sup>115</sup> Specificity and affinity studies in lectin/carbohydrate interaction. Sulak O., lameignere E., Wimmerova M., Imberty A. *Carbohydr Chem.* **2009**, *35*, 356-371.
- <sup>116</sup> Lectins: tools for the molecular understanding of the glycode. Ambrosi M., Cameron N. and Davis B. *Org Biomol Chem.* **2005**, *3*, 1593-1608.
- <sup>117</sup> Binding characteristics of N-acetylglucosamine-specific lectin of the isolated chicken hepatocytes: similarities to mammalian hepatic galactose/N-acetylgalactosamine-specific lectin. Lee R., Rice K., Rao N., Ichikawa Y., Barthel T., Piskarev V. and Lee Y. *Biochemistry* **1989**, *28*(21), 8351-8358.
- <sup>118</sup> Affinity enhancement by multivalent lectin-carbohydrate interaction. Lee R and Lee Y. *Glycoconjugate J* **2000**, *17*(7-9), 543-551.
- <sup>119</sup> Cell-surface carbohydrate recognition by animal and viral lectins. Weis W. *Curr. Opin. Struct. Biol.* *7*, 624 (1997).
- <sup>120</sup> Structural basis of lectin-carbohydrate recognition. Weis W. and Drickamer K. *Annu Rev Biochem.* **1996**, *65*, 441-473.
- <sup>121</sup> A unified vision of the building blocks of life. Marth J. *Nat Cell Biol.* **2008**, *10*, 1015-1016.

- 
- <sup>122</sup> Zur Kenntnis der antifermentativen, lytischen und agglutinierenden Wirkungen des Blutsersums und der Lymphe. Landsteiner K. *Zentralbl Bakteriol.* **1900**, *27*, 357-362.
- <sup>123</sup> Immunochemical studies on blood groups. XVII. structural units involved in blood group A and B specificity. Kabat E. and Leskowitz S. *J Am Chem Soc.* **1955**, *77*, 5159-5164.
- <sup>124</sup> The inhibition of the haemagglutinins in plant seeds by human blood group substances and simple sugars. Morgan W. and Watkins W. *Br J Exp Pathol*, **1953**, *34*, 94-103.
- <sup>125</sup> Inhibition by simple sugars of enzymes which decompose the blood-group substances. Morgan W. and Watkins W. *Nature* **1955**, *175*, 676-677.
- <sup>126</sup> Glycotherapy: New Advances Inspire a Reemergence of Glycans in Medicine. Hudak J. and Bertozzi C. *Chem Biol.* **2014**, *21*(1), 16-37.
- <sup>127</sup> The elaboration of specific soluble substance by Pneumococcus during growth. Dochez A. and Avery O. *J Exp Med.* **1917**, *26*, 477-493.
- <sup>128</sup> Cutaneous reactions in pneumonia: the development of antibodies following the intradermal injection of type-specific polysaccharide. Francis T. and Tillett W. *J Exp Med.* **1930**, *52*, 573-585.
- <sup>129</sup> Persistence of antibodies in human subjects injected with pneumococcal polysaccharides. Heidelberger M., Dilapi M. and Siegel W. *J Immunol.* **1950**, *65*, 535-541.
- <sup>130</sup> Vaccines in the era of genomics: the pneumococcal challenge. Barocchi M., Censini S. and Rappuoli R. *Vaccine* **2007**, *25*, 2963-2973.
- <sup>131</sup> Carbohydrate based vaccines. Kuberan B. and Linhardt R. *Curr Org Chem.* **2000**, *4*, 653-677.
- <sup>132</sup> From the laboratory to the clinic: A retrospective on fully synthetic carbohydrate-based anticancer vaccines. Danishefsky S. and Allen J. *Angew Chem Int Ed Engl* **2000**, *39*, 836-863.
- <sup>133</sup> Towards the development of antitumor vaccines: A synthetic conjugate of a tumor-associated MUC1 glycopeptide antigen and a tetanus toxin epitope. Keil S., Claus C., Dippold W. and Kunz H. *Angew Chem Int Ed Engl.* **2001**, *40*, 366-369.

- 
- <sup>134</sup> Carbohydrates and Immunology: Synthetic Oligosaccharide Antigens for Vaccine Formulation. Morelli L., Poletti L., Lay L. *Eur J Org Chem.* **2011**, *29*, 5723-5777.
- <sup>135</sup> Synthesis and medical applications of oligosaccharides. Seeberger P and Werz D. *Nature* **2007**, *446*, 1046-1051.
- <sup>136</sup> A synthetic conjugate polysaccharide vaccine against Haemophilus influenzae type b. Verez-Bencomo V., Fernández-Santana V., Hardy E., Toledo M., Rodríguez M., Heynngnezz L., Rodriguez A., Baly A., Herrera L., Izquierdo M. *et al. Science* **2004**, *305*, 522-525.
- <sup>137</sup> Rapid synthesis of a glycosylphosphatidylinositol-based malaria vaccine using automated solid-phase oligosaccharide synthesis. Hewitt M., Snyder D and Seeberger P. *J Am Chem Soc.* **2002**, *124*, 13434-13436.
- <sup>138</sup> Synthetic GPI as a candidate anti-toxic vaccine in a model of malaria. Schofield L., Hewitt M., Evans K., Siomos M. and Seeberger P. *Nature* **2002**, *418*, 785-789.
- <sup>139</sup> Insights in the rational design of synthetic multivalent glycoconjugates as lectin ligands. Deniaud D., Julienne K. and Gouin S. *J Org Biomol Chem.* **2011**, *9*, 966-979.
- <sup>140</sup> Aminoglycoside antibiotics in the 21st century. Becker B and Cooper M. *ACS Chem Biol.* **2013**, *8*, 105-115.
- <sup>141</sup> Recent trends in glycodendrimer syntheses and applications. Chabre Y. and Roy R. *Curr Top Med Chem.* **2008**, *8*, 1237-1285.
- <sup>142</sup> Glyconanoparticles as multifunctional and multimodal carbohydrate systems. Marradi M., Chiodo F., García I. and Penadés S. *Chem Soc Rev.* **2013**, *42*, 4728-4745.
- <sup>143</sup> Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands. Kitov P., Sadowska J., Mulvey G., Armstrong G., Ling H., Pannu N., Read R. and Bundle D. *Nature* **2000**, *403*, 669-672.
- <sup>144</sup> Dendrimer-like PEO glycopolymers exhibit anti-inflammatory properties. Rele S., Cui W., Wang L., Hou S., Barr-Zarse G., Tatton D., Gnanou Y., Esko J., Chaikof L. *J Am Chem Soc.* **2005**, *127*, 10132-10133.

- 
- <sup>145</sup> Influencing receptor-ligand binding mechanisms with multivalent ligand architecture. Gestwicki J., Cairo C., Strong L., Oetjen K. and Kiessling L. *J Am Chem Soc.* **2002**, *124*, 14922-14933.
- <sup>146</sup> PNA–sugar conjugates as tools for the spatial screening of carbohydrate–lectin interactions. Scheibe C. and Seitz O. *Pure Appl Chem.* **2012**, *84*(1), 77-85.
- <sup>147</sup> Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. Merrifield R. *J Am Chem Soc.* **1963**, *85*(14), 2149-2154.
- <sup>148</sup> Circular dichroism and the conformational analysis of bio-molecules. Fasman G. *Plenum Press*, New York, **1996**.
- <sup>149</sup> Microscopy. II. Electron microscopy: A review. Selby C. *Cancer Res.* **1953**, *13*(11), 753-775.
- <sup>150</sup> Negative staining of proteins. Kiselev N., Sherman M. and Tsuprun V. *Electron Microsc Rev.* **1990**, *3*(1), 43-72.
- <sup>151</sup> Phase contrast electron microscopy: development of thin-film phase plates and biological applications. Nagayama K. and Danev R. *Phil Trans R Soc. B* **2008**, *363*, 2153-2162.
- <sup>152</sup> Flow Cytometry: Principles and Clinical Applications in Hematology. Brown M. and Wittwer C. *Clinic Chem.* **2000**, *46*(8), 1221-1229.
- <sup>153</sup> [http://docs.abcam.com/pdf/protocols/Introduction\\_to\\_flow\\_cytometry\\_May\\_10.pdf](http://docs.abcam.com/pdf/protocols/Introduction_to_flow_cytometry_May_10.pdf)
- <sup>154</sup> The ELISA Guidebook. Walker J. Springer Protocols, *Methods in Molecular Biology* 516, Humana Press, **2009**.
- <sup>155</sup> Cloning and screening of sequences expressed in a mouse colon tumor. Augenlicht L., Kobrin D. *Cancer Res.* **1982**, *42*(3), 1088-93.
- <sup>156</sup> Identification of interferon-modulated proliferation-related cDNA sequences. Kulesh D., Clive D., Zarlenga D. and Greene J. *Proc Natl Acad Sci USA* **1987**, *84*(23), 8453-8457.
- <sup>157</sup> Glycan array: a powerful tool for glycomics studies. Liang C. and Wu C. *Expert Rev Proteomics* **2009**, *6*(6), 631-645.

- 
- <sup>158</sup> The Asialoglycoprotein Receptor: Relationships Between Structure, Function, and Expression. Stockert R. *Physiol Rev.* **1995**, *75*, 591-609.
- <sup>159</sup> Detection of the asialoglycoprotein receptor on cell lines of extrahepatic origin. Park J., Cho E., Shin S., Lee Y. and Kim K. *Biochem Biophys Res Commun* **1998**, *244*, 304-11.
- <sup>160</sup> Targeting C-type lectin receptors with multivalent carbohydrate ligands. Lepenies B., Leec J. and Sonkaria S. *Adv Drug Deliv Rev.*, **2013**, *65*, 1271–1281.
- <sup>161</sup> The role of Carbohydrate in the Hepatic Recognition and Transport of Circulating Glycoprotein. Ashwell G. and Morell A. *Adv Enzymol Relat Areas Mol Biol.* **1974**, *41*, 99-128.
- <sup>162</sup> Binding of synthetic oligosaccharides to the hepatic Gal/GalNAc lectin. Dependence on fine structural features. Lee Y., Townsend R., Hardy M., Lönnngren J., Arnarp J., Haraldsson M. and Lönn H. *J Biol Chem.* **1983**, *258*(1), 199-202.
- <sup>163</sup> Major and Minor Forms of the Rat Liver Asialoglycoprotein Receptor are Independent Galactose-Binding Proteins. Primary Structure and Glycosylation Heterogeneity of Minor Receptor Forms. Halberg D., Wager R., Farrell D., Hildreth J., Quesenberry M., Loeb J., Holland E. and Drickamer K. *J Biol Chem.* **1987**, *262*, 9828-9838.
- <sup>164</sup> Crystal Structure of the Carbohydrate Recognition Domain of the H1 Subunit of the Asialoglycoprotein Receptor. Meier M., Bider M., Malashkevich V., Spiess M. and Burkhard P. *J Mol Biol.* **2000**, *300*, 857-865.
- <sup>165</sup> Interactions of Oligosaccharides and Glycopeptides with Hepatic Carbohydrate Receptors. Lee Y., Lee R., Ernst B., Hart G. and Sinaý P. *Carbohydrates in Chemistry and Biology*: Wiley-VCH Verlag GmbH. **2000**, 549-561.
- <sup>166</sup> Trivalent, Gal/GalNAc-Containing Ligands Designed for the Asialoglycoprotein Receptor. Khorev O., Stokmaier D., Schwardt O., Cutting B. and Ernst B. *Bioorg Med Chem.* **2008**, *16*, 5216-5231.
- <sup>167</sup> Preparation of a high-affinity photolabeling reagent for the Gal/GalNAc lectin of mammalian liver: demonstration of galactose-combining sites on each subunit of rabbit hepatic lectin. Lee R. and Lee Y. *Biochemistry* **1986**, *25*, 6835–6841.

- 
- <sup>168</sup> Modification of triantennary glycopeptide into probes for the asialoglycoprotein receptor of hepatocytes. Rice K. and Lee Y. *J Biol Chem.* **1990**, *265*, 18423–18428.
- <sup>169</sup> Lactoferrin binding to the rat asialoglycoprotein receptor requires the receptor's lectin properties. McAbee D., Jiang X and Walsh K. *Biochem J.* **2000**, *348*(Pt 1), 113-117.
- <sup>170</sup> Mechanism of N-Acetylgalactosamine Binding to a C-type Animal Lectin Carbohydrate-recognition Domain. Kolatkar A., Leung A., Isecke R., Brossmer R., Drickamer K. and Weis W. *J Biol Chem.* **1998**, *273*, 19502-8.
- <sup>171</sup> The asialoglycoprotein receptor: a model for endocytic transport receptors. Spiess M. *Biochemistry* **1990**, *29*(43), 10009-10018.
- <sup>172</sup> Essential Cell Biology, IV edition. Jonson A., Lewis J., Walter P., Roberts K., Raff M., Hopkin K., Bray D. and Alberts B. *Protein Structure and Function* **2013**.
- <sup>173</sup> Studies of helix fraying and solvation using <sup>13</sup>C' isotopomers. Fesinmeyer M., Peterson E., Dyer B. and Andersen N. *Protein Sci.* **2005**, *14*(9), 2324-2332.
- <sup>174</sup> Thermodynamics of Multivalent Interactions: Influence of the Linker. Kane R. *Langmuir* **2010**, *26*(11), 8636-8640.
- <sup>175</sup> 'Non-specific' binding. The problem, and a solution. Mendel C. and Mendel D. *Biochem J.* **1985**, *228*(1), 269-272.
- <sup>176</sup> Identification of amino acid residues that determine pH dependence of ligand binding to the asialoglycoprotein receptor during endocytosis. Wragg S. and Drickamer K. *J Biol Chem.* **1999**, *274*(50), 35400-35406.
- <sup>177</sup> The Actin Cytoskeleton Is Required for Receptor-mediated Endocytosis in Mammalian Cells. Lamaze C., Fujimoto M., Yin H. and Schmid S. *J Biol Chem.* **1997**, *272*, 20332-20335.
- <sup>178</sup> Tailored Presentation of Carbohydrates on a Coiled Coil-Based Scaffold for Asialoglycoprotein Receptor Targeting. Zacco E., Hütter J., Heier J., Mortier J., Seeberger P., Lepenies B. and Kokschi B. *ACS Chem Biol.* **2015**, DOI: 10.1021/acscchembio.5b00435.
- <sup>179</sup> The Kinetics of the Removal of the N-Methyltrityl (Mtt) Group During the Synthesis of Branched Peptides. Li D. and Elbert D. *J. Pept. Res.* **2002**, *60*, 300-303.

- 
- <sup>180</sup> Principles of vaccine design—Lessons from nature. Zepp F. *Vaccine* **2010**, 28S, C14-C24.
- <sup>181</sup> Carbohydrate based vaccines. Vliegthart J. *FEBS Letters* **2006**, 580, 2945-2950.
- <sup>182</sup> Toward automated synthesis of oligosaccharides and glycoproteins. Sears P. and Wong C. *Science* **2001**, 291(5512), 2344-2350.
- <sup>183</sup> Recent advances in automated solid-phase carbohydrate synthesis: from screening to vaccines. Plante O. and Seeberger P. *Curr Opin Drug Discov Devel* **2003**, 6(4), 521-525.
- <sup>184</sup> Naturally occurring antibodies directed against carbohydrate tumor antigens. Schwartz-Albiez R. *Adv Exp Med Biol.* **2012**, 750, 27-43.
- <sup>185</sup> Monoclonal Anti-idiotypic Antibodies against Carbohydrate-associate Epitope for Anti-Cancer Vaccine Development. Lee G., Cheung A., Ge B., Zhu M., Li P., Hsu E. and Huang K. *J Vaccines Vaccin* **2010**, 1, 106-112.
- <sup>186</sup> Epigenetic regulation of glycosylation is the quantum mechanics of biology. Lauca G., Vojtác A. and Zoldoš V. *Bioch Bioph Acta - General Subjects* 2014, 1840(1), 65-70.
- <sup>187</sup> The promise of glycomics, glycan arrays and carbohydrate-based vaccines. Lepenies B. and Seeberger P. *Immunopharmacol Immunotoxicol*, **2010**, 32(2), 196-207.
- <sup>188</sup> Recent developments in synthetic oligosaccharidebased bacterial vaccines. Pozsgay V. *Curr Top Med Chem* **2008**, 8(2), 126-240.
- <sup>189</sup> Synthetic 6B di-, tri-, and tetrasaccharide-protein conjugates contain pneumococcal type 6A and 6B common and 6B-specific epitopes that elicit protective antibodies in mice. Jansen W., Hogenboom S., Thijssen M., Kamerling J., Vliegthart J., Verhoef J., Snippe H., Verheul A. *Infect Immun* **2001**, 69(2): 787-93.
- <sup>190</sup> Vaccination against typhoid fever: present status. Ivanoff B., Levine M. and Lambert P. *Bull World Health Organ.* **1994**, 72, 957–791.
- <sup>191</sup> Synthesis and medical applications of oligosaccharides. Seeberger P. and Werz D. *Nature* **2007**, 446(7139), 1046-1051.



- 
- <sup>192</sup> Potential recombinant vaccine against influenza A virus based on M2e displayed on nodaviral capsid nanoparticles. Yong C., Yeap S., Ho K., Omar A. and Tan W. *Int J Nanomedicine* **2015**, *10*, 2751-2763.
- <sup>193</sup> Rational Engineering of Recombinant Picornavirus Capsids to Produce Safe, Protective Vaccine Antigen. Porta1C., Kotecha A., Burman A., Jackson T., Ren J., Loureiro S., Jones I., Fry E., Stuart D. and Charleston B. *PLoS Pathog* **2013**, *9*(3), e1003255.
- <sup>194</sup> A review of clinical trials of human papillomavirus prophylactic vaccines. Schiller J., Castellsague X., Garland S. *Vaccine* **2012**, *30*(15), F123-138.
- <sup>195</sup> MUC1 as a target antigen for cancer immunotherapy. Acres B. and Limacher J. *Expert Rev Vaccines*. **2005**, *4*(4), 493-502.
- <sup>196</sup> Automated Synthesis of the Tumor-Associated Carbohydrate Antigens Gb-3 and Globo-H: Incorporation of r-Galactosidic Linkages. Werz D., Castagner B. and Seeberger P. *J Am Chem Soc*. **2007**, *129*, 2770-2771
- <sup>197</sup> Ligation of B and T lymphocyte attenuator prevents the genesis of experimental cerebral malaria. Lepenies B., Pfeffer K., Hurchla M., Murphy T., Murphy K., Oetzel J, Fleischer B. and Jacobs T. *J Immunol* **2007**, *179*(6), 4093-4100.
- <sup>198</sup> Enhancement of the immunogenicity of synthetic carbohydrates by conjugation to virosomes: a leishmaniasis vaccine candidate. Liu X., Siegrist S., Amacker M., Zurbriggen R., Pluschke G. and Seeberger P. *ACS Chem Biol* **2006**, *1*(3), 161-164.
- <sup>199</sup> Functional aspects of *Leishmania donovani* lipophosphoglycan during macrophage infection. Descoteaux A. and Turco S. *Microbes Infect.* **2002**, *4*, 975- 981.
- <sup>200</sup> Immunogenicity and Diagnostic Potential of Synthetic Antigenic Cell Surface Glycans of *Leishmania*. Anish C., Martin C., Wahlbrink A., Bogdan C, Ntais P, Antoniou M. and Seeberger P. *ACS Chem. Biol.*, **2013**, *8*(11), 2412-2422.
- <sup>201</sup> *Leishmania chagasi*: lipophosphoglycan characterization and binding to the midgut of the sand fly vector *Lutzomyia longipalpis*. *Mol Biochem Parasitol.* **2002**, *121*, 213-224.

---

<sup>202</sup> Major Histocompatibility Complex: Interaction with Peptides. Liu J. and Gao G. *eLS* **2011**. DOI: 10.1002/9780470015902.a0000922.pub2.

<sup>203</sup> Synthetically defined glycoprotein vaccines: current status and future directions. Adamo R., Nilo A., Castagner B., Boutureira O., Bertia F. and Bernardes G. *Chem Sci.* **2013**, *4*, 2995-3008.

<sup>204</sup> Structural basis for lack of toxicity of the diphtheria toxin mutant CRM197. Malito E., Bursulaya B., Chen C., Lo Surdo P., Picchianti M., Balducci E., Biancucci M., Brock A., Berti F., Bottomley M., Nissum M., Costantino P., Rappuoli R. and Spraggon G. *Proc Natl Acad Sci USA* **2012**, *109*(14), 5229-5234.

<sup>205</sup> Epitope Density Influences CD8+ Memory T Cell Differentiation. Leignadier J and Labrecque N. *PLoS ONE* **2010**, *5*(10), e13740.

<sup>206</sup> Immunogenicity of two accelerated hepatitis B vaccination protocols in liver transplant candidates. Engler S., Sauer P., Golling M., Klar E., Benz C., Stremmel W. and Kallinowski B. *Eur J Gastroenterol Hepatol.* **2001**, *13*(4), 363-367.

<sup>207</sup> New CD4+ and CD8+ T cell responses induced in chronically HIV type-1-infected patients after immunizations with an HIV type 1 lipopeptide vaccine. Gahery H., Daniel N., Charmeteau B., Ourth L., Jackson A., Andrieu M., Choppin J., Salmon D., Pialoux G. and Guillet J. *AIDS Res Hum Retroviruses.* **2006**, *22*(7), 684-694.

<sup>208</sup> Epitopes for human CD4+ cells on diphtheria toxin: structural features of sequence segments forming epitopes recognized by most subjects. Raju R., Navaneetham D., Okita D., Diethelm-Okita B., McCormick D., Conti-Fine B. *Eur J Immunol.* **1995**, *25*(12), 3207-3214.

<sup>209</sup> T-cell epitope as carriers molecule for conjugate vaccines. Bixler G., Pillai S. and Insel R. Patent **1989**, WO 1989006974 A2.

<sup>210</sup> Synthetic peptide representing a T-cell epitope of CRM197 substitutes as carrier molecule in a Haemophilus influenzae type B (Hib) conjugate vaccine. Bixler G., Eby R., Dermody K., Woods R., Seid R. and Pillai S. *Adv Exp Med Biol.* **1989**, *251*, 175-180.

- 
- <sup>211</sup> Quantitative Investigation of the Modular Primer Effect for DNA and Peptide Nucleic Acid Hexamers. Nilsson P., O'meara D., Edebratt F., Persson B., Uhlén M., Lundeberg J. and Nygren P. *Anal Biochem.* **1999**, *269*, 155-161.
- <sup>212</sup> The Use of a Neutral Peptide Aptamer Scaffold to Anchor BH3 Peptides Constitutes a Viable Approach to Studying Their Function. Stadler L., Tomlinson D., Lee T., Knowles M. and Ko Ferrigno P. *Cell Death Dis.* **2014**, *5*, e1037.
- <sup>213</sup> Carrier-induced epitopic suppression, a major issue for future synthetic vaccines. Schutze M., Leclerc C., Jolivet M., Audibert F. and Chedid L. *J Immunol.* **1985**, *135*(4), 2319-2322.
- <sup>214</sup> Oligosaccharide-protein conjugates as vaccine candidates against bacteria. Pozsgay V. *Adv Carbohydr Chem Biochem.* **2000**, *56*, 153-199.
- <sup>215</sup> Vaccines based on the cell surface carbohydrates of pathogenic bacteria. Jones C. *An Acad Bras Cie<sup>nc</sup>* **2010**, *77*, 293-324.
- <sup>216</sup> Carbohydrate moieties as vaccine candidates. Lucas A., Apicella M. and Taylor C. *Clin Inf Diseases* **2005**, *41*, 705-712.
- <sup>217</sup> Neutralizing Antiviral B Cell Responses. Bachmann M., Zinkernagel R. *Annu. Rev. Immunol.* **1997**, *15*, 235–270.
- <sup>218</sup> Pathogen Recognition and Inflammatory Signaling in Innate Immune Defenses. Mogensen T. *Clin. Microbiol. Rev.* **2009**, *22*, 240–273.
- <sup>219</sup> . In Situ Recognition of Cell-Surface Glycans and Targeted Imaging of Cancer Cells. Xu X., Cheng H., Chen W., Cheng S., Zhuo R. and Zhang X. *Sci Rep.* **2013**, *3*, 2679.
- <sup>220</sup> MHC Class I Presented T Cell Epitopes as Potential Antigens for Therapeutic Vaccine against HBV Chronic Infection. Comber J., Karabudak A., Shetty V., Testa J., Huang X. and Philip R. *Hepat Res Treat.* **2014**, ID:860562.
- <sup>221</sup> Inhibition der Bildung amyloider Aggregate. *Brandenburg, E. DDC 540 Chemie*, **2012**.
- <sup>222</sup> Carbohydrate-based vaccines. Kuberan B. and Linhardt R. *Curr Org Chem.* **2000**, *4*(6), 653-677.

- 
- <sup>223</sup> A de novo peptide hexamer with a mutable channel. Zaccai N., Chi B., Thomson A., Boyle A., Bartlett G., Bruning M., Linden N., Sessions R., Booth P., Brady R. and Woolfson D. *Nat Chem Biol.* **2011**, 7(12):935-41.
- <sup>224</sup> Assembly pathway of a designed alpha-helical protein fiber. Bromley E., Channon K., King P., Mahmoud Z., Banwell E., Butler M., Crump M., Dafforn T., Hicks M., Hirst J., Rodger A. and Woolfson D. *Biophys J.* **2010**, 98, 1668-1676.
- <sup>225</sup> Refined structure of monomeric diphtheria toxin at 2.3 Å resolution. Bennett M. and Eisenberg D. *Protein Sci.* **1994**, 3, 1464. PDB entry: 1MDT.
- <sup>226</sup> The Beta-Turn Conformation in Wheat Gluten Proteins: Relationship to Gluten Elasticity. Tatham A., Mifflin B. and Shewry P. *Cereal Chem.* **1985**, 62, 405-412.
- <sup>227</sup> The prolamin storage proteins of cereal seeds: structure and evolution. Shewry P. and Tatham A. *Biochem J.* **1990**, 267, 1-12.
- <sup>228</sup> Assembly Pathway of a Designed  $\alpha$ -Helical Protein Fiber. Bromley E., Channon K., King P., Mahmoud Z., Banwell E., Butler M., Crump M., Dafforn T., Hicks M., Hirst J., Rodger A. and Woolfson D. *Biophys. J.* **2010**, 98, 1668-76.
- <sup>229</sup> Assessment of IgG Avidity Against Pertussis Toxin and Filamentous Hemagglutinin Via an Adapted Enzyme-Linked Immunosorbent Assay (ELISA) Using Ammonium Thiocyanate. Almanzar G., Ottensmeier B., Liese J., Prelog M. *J Immunol Methods* **2013**, 387, 36-42.
- <sup>230</sup> Fluoresceinated lectins as probe for cell surface changes associated with lymphocyte transformation. Strauchen J. *Am J Hematol.* **1982**, 12, 227-232.
- <sup>231</sup> Lectins as pattern recognition molecules: The effects of epitope density in innate immunity. Dam T. and Brewer C. *Glycobiology* **2010**, 3, 270-279.
- <sup>232</sup> A Self-Assembling Peptide Scaffold for the Multivalent Presentation of Antigens. Zacco E., Anish C., Martin C., von Berlepsch H., Brandenburg E., Seeberger P. and Kokschi B. *Biomacromolecules* **2015** [Epub ahead of print].
- <sup>233</sup> How to measure and predict the molar absorption coefficient of a protein. Pace C., Vajdos F., Fee L., Grimsley G. and Gray T. *Protein Sci.* **1995**, 4, 2411-2423.

---

# 7

**Synthetic Glycopeptides as Multivalent Scaffolds for Carbohydrates:**  
from receptor targeting to vaccines exploiting sugar-protein interactions

## CURRICULUM VITAE

---



---

For reasons of data protection, the curriculum vitae is not published in the electronic version.

---

For reasons of data protection, the curriculum vitae is not published in the electronic version.



