UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE FARMACIA

Departamento de Bioquímica y Biología Molecular II

PATRONES DE GLICOSILACIÓN

Y

RECONOCIMIENTO POR LECTINAS

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

Ioanna Kalograiaki

Directora: Dra. María Dolores Solís Sánchez

Tutora: Prof. Pilar Iniesta Serrano

Madrid, 2017

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE FARMACIA

Departamento de Bioquímica y Biología Molecular II

GLYCOSYLATION PATTERNS

AND

RECOGNITION BY LECTINS

PhD Dissertation presented by
Ioanna Kalograiaki

Supervisor: Dr. María Dolores Solís Sánchez

Tutor: Prof. Pilar Iniesta Serrano

Madrid, 2017

ACKNOWLEDGEMENTS (I)

The present PhD Thesis has been accomplished at the Group of Protein Structure and Dynamics, Department of Biological Physical Chemistry, Rocasolano Institute of Physical Chemistry of the Spanish National Research Council (IQFR-CSIC, Madrid, Spain) under the supervision of Dr. Dolores Solís.

Part of the experimental work was performed at the Department of Animal Health, Institute of Agrobiotechnology, Spanish National Research Council-Universidad Pública de Navarra-Government of Navarra (Pamplona, Spain) in collaboration with the group of Dr. Junkal Garmendia García, during a short-term stay funded by CIBERES. Also, thanks to the Marie Curie Fellowship, short-stays in other laboratories across Europe, contributed to the training of the PhD candidate. In detail: at the Department of Genetics, Cell- and Immunobiology of the Semmelweis University (Budapest, Hungary) in collaboration with the group of Prof. Edit I. Buzás; at the Department of Organic Chemistry of the Royal Institute of Technology (KTH, Stockholm, Sweden) in collaboration with the group of Prof. Olof Ramström; at Attana AB (Stockholm, Sweden) under the supervision of Dr. Teodor Aastrup. Finally, the PhD candidate was recruited in the Nuclear Magnetic Resonance and Molecular Recognition group, department of Chemical & Physical Biology of the Biological Research Center, Spanish National Research Council (CIB-CSIC, Madrid, Spain), where she has obtained part of the results here presented, under the supervision of Prof. F. Javier Cañada.

Ioanna Kalograiaki is a Marie Curie Fellow and was funded by an Early Stage Researcher contract from the European Commission in the context of the Marie Curie Initial Training Network DYNANO (Grant PITN-GA-2011-289033). Financial support from the Spanish Ministry of Economy and Competitiveness (Grants BFU2012-36825, BFU2015-70052-R), the CIBER of Respiratory Diseases (CIBERES), an initiative from the Spanish Institute of Health Carlos III (ISCIII), and the Marie Curie ITN GLYCOPHARM (Grant PITN-GA-2012-317297) is also gratefully acknowledged.

Στην οικογένειά μου και στη μνήμη του Παππού

A mi familia y en memoria de mi Abuelo

ACKNOWLEDGEMENTS (II)

First and foremost, I would like to express my most sincere gratitude to my supervisor, Dr. Dolores Solís, for giving me the opportunity to enrol in this PhD. She opened a whole new world beyond my eyes once suggesting me to dive into the fast-paced, evolving field of Glycosciences. She will always have my respect and appreciation for inspiring my research through stimulating conversations, for always being open to ground-breaking ideas and scientific "adventures", as well as for offering me all the opportunities I could have ever asked for, to grow as a research scientist both inside and outside the lab. Thank you, Lola, for all your confidence, your support in every step, your contagious passion for discovery and the persistence to carry out the research the way it should be done.

I would also like to warmly thank Dr. Margarita Menéndez for being my mentor throughout these four years at the Rocasolano Institute of Physical Chemistry; for her valuable time, endless advice, for teaching me methods, concepts, good scientific practices and, above all, for keeping her door always open. I am thankful to Dr. Maríasun Campanero-Rhodes for introducing me to the designer´ microarray technology and infinite capacity, and always sharing her "fresh" ideas. Correspondingly, I feel indebted to my current supervisor at the Biological Research Center, Prof. F. Javier Cañada, for his continuous support and patience in teaching me NMR, for sharing his passion for quality science and, of course, his infinite innovative spirit. Words are not enough to express my appreciation for the prospects and motivation you inculcated on me to further learn and evolve in the glycoscience field.

At this point, I would like to thank the European Commission and Research Executive Agency for funding the Marie Curie Actions Initial Training Network DYNANO and offering me an Early Stage Researcher fellowship that lasted three years and financed a big part of this Thesis. In this context, my gratitude is colossal for our collaborators who warmly welcomed me in their labs during my secondments. Starting from the "nationals", I greatly appreciate the invitation of Dr. Junkal Garmendia for visiting her lab and her enthusiastic contribution to my training; Dr. Begoña Euba for her constant efforts, the excellent teamwork and cooperation. Moving on to our european partners, I would like to specially thank Dr. Teodor Aastrup for opening up the doors of his enterprise for me and for securing my access to cutting-edge equipment and, Dr. Davide Proverbio for his ever valuable help as well as for the personal training sessions. I would also like to express my deep appreciation to Profs. Edit Buzás and Olof Ramström for allowing me to participate in motivating and multicultural scientific environments. Last but not least, I would like to thank Prof. Stéphane Vincent and Dr. Marta Abellán-Flos for the smooth, first-rate collaboration, Dr. Laura Lagartera for all the support and training on SPR, as well as Profs. Ernesto and Pedro García and Dr. José Manuel Andreu for permitting me the access to their lab facilities and instrumentation.

Moving on, I would like to wholeheartedly thank my lab mates, Lara, Begoña, Mónica and Radek for the incessant backing and long-term friendship. Lara, thanks for being my older sister here in Spain, always there for me, sharing precious moments. Also, Javito and Cris, Miguel Mompeán, Manu, María, Erney, Tere, Mon, Rosa, Palma, Javier, Iván, Noe and Mer, thank you for the enriching coexistence at the Institute and all the moments we shared chilling out. Similarly, this time as a CIB fellow, I would like to warmly thank Ángeles, Carmen, Beatriz, Javierín, Anita and Borja for making my integration in the new team almost instant and for the great environment maintained at work.

From the Faculty of Pharmacy of the Complutense University, I would like to express my honest gratitude to my tutor, Prof. Pilar Iniesta, for the time invested and all the support, both academic and emotional. Also, to Prof. Juan Torrado for being my supervisor during the BSc and MSc Theses and the person who encouraged me to "take the step" and advance in research.

On a more personal note, one step away from this PhD completion I could never forget the contributions of people that came into my life recently but have thoroughly supported me along the way: Elli, Dimi, Estefanía, Elena, Isa, Phoebus and Vassilis, my most sincere thank you. To Alessandra, my warmest hug and infinite gratitude for her unreserved friendship and always optimistic attitude, sharing this hard way to success.

I will be eternally grateful to my family back in Greece, the people I mostly cherish and miss every day. They have always supported my decisions, let me follow my dreams, taught me the values of life and the hard work, to rely on my own strength, to set goals and accomplish them. No words can express my thankfulness for their unconditional love, their continuous sacrifice, the dedication and their respect to my personal freedom, and no actions can pay these back. Mom, dad, all I am, I owe it to you. A huge part of this Thesis, too. To my sister Chara, special thanks for being one of the best friends I could ever have, my English teacher and text editor, and an inimitable pillar in life; to my brother Stefanos for his nonstop cheering up and love; to my aunt, Arísti, for inculcating on me her always positive thinking and, finally, to my grandparents, the one who is still with us and the ones who passed away, wonderful persons that have been the example to follow for all of us.

Back in time, I would like to acknowledge my first Teacher, Spyros Monastiriotis, for being the first to believe in me and motivate me to study a carreer in Science, teaching me how to state questions and design the strategy that leads to the (most possible) solution, and having always followed my progress thereafter.

To all of you, thank you from the bottom of my heart,

Ioanna

I am indebted to my father for living, but to my teacher for living well.

Alexander the Great, 356-323 B.C.

TABLE OF CONTENTS

ABSTRACT

GLYCOSYLATION PATTERNS AND RECOGNITION BY LECTINS

INTRODUCTION

The space on the cell surface is fairly limited and has to harbour a multitude of signals, making high-density coding compulsory. Satisfying this requirement, the glycocalyx, a fuzz-like carbohydrate coat, is present on the external surface of cell membranes of both eukaryotes and prokaryotes, as well as of cell-derived vesicles. In addition to offering chemical resistance and backbone support, their extended presence on cellular surfaces plays an important functional role, frequently mediating recognition processes. Lectins, a ubiquitous family of effector proteins, are capable of reversibly binding carbohydrate sequences with extraordinary selectivity, thus acting as translators of sugar-encoded signals.

Their existence and action as fine-tuned binders argues for their probable co-evolution with saccharides in the frame of the high organisational diversity of the latter. Through recognition of sugar structures by endogenous lectins, the discrimination between self- and non-self-glycosylation patterns lies behind the innate immunity response. In this perspective, the eventual presentation of aberrant glycosylation profiles by host cells at both structural and expression levels is commonly associated to pathogenesis, cancer being the most relevant example.

Even though these altered self-patterns do not typically elicit immune responses, they constitute attractive targets for diagnosis and therapeutic intervention. During the last decade, the findings supporting that extracellular vesicles (EV) could transfer molecular patterns, metabolites or effectors in a horizontal manner, revolutionised the biomarker field. These membrane vesicles are assembled in an orchestrated process, then liberated by almost all cells, including tumour cells, and are thought to share glycosylation profiles with them. Although the use of lectins as tools for glycotypification of tissues in physiology, or of bacterial strains in microbiology, was long established, lectin-based high-scale analytical methodologies for classification of EV are still missing.

In contrast, nonself pathogens´ glycosignatures, with definite structure, accessibility and density, are key targets for recognition by host lectins and/or soluble molecules for triggering defence responses. Nonetheless, bacteria are highly dynamic organisms and effortlessly adapt to changing environments under selection pressure. In order to evade

host immunity strategies, they normally disguise their tell-tale glycostructures or even integrate host mimicry motifs. Along with pathogen glycotypification using lectins of known specificity as decoding tools, it is expected that the exploration of docking sites for endogenous lectins directly at the bacterial surface will pave the way for the establishment of virulence correlations. Up until recently, these studies were commonly performed following isolation and purification of selected surface components. However, in these ligand-targeted approaches, the natural presentation and accessibility of glycans along with any possible contribution to the binding of neighbour molecules are not taken into account, an aspect of particular relevance in the case of weak binders.

OBJECTIVES

The core objective of this PhD Thesis is the design and development of novel high-throughput platforms for the study of protein-carbohydrate interactions directly at the surface of the probes. Specific objectives are the following:

- Design and application of novel microarray and Quartz Crystal Microgravimetry strategies for the analysis of surface glycosylation of bacteria and the exploration of their recognition by lectins of the innate immune system. Nontypeable *Haemophilus influenzae* (NTHi) was selected as model of non-encapsulated opportunistic pathogen of relevance in upper respiratory tract infections.

- Development of microarray and QCM chips based on the incorporation of entire extracellular vesicles, as powerful tools for the analysis of glycosylation patterns of exosomes and microvesicles of diverse cellular origin and determination of the kinetic parameters of vesicle-lectin interactions.

- Development of a microarray platform for the evaluation of adhesin functionality in live bacteria and the screening of glycosylated nano-scaffolds as possible anti-adhesive agents.

FUNDAMENTAL INPUTS OF THE PHD THESIS

In this Thesis, novel microarray and quartz crystal microbalance (QCM) strategies were applied to the analysis of carbohydrate-mediated interactions directly at the bacterial or extracellular vesicle surface, thus preserving the native environment, distribution and vicinity of glycan targets.

First, a novel bacteria-based microarray set-up was developed and the glycosylation patterns of different strains of the opportunistic respiratory pathogen nontypeable *Haemophilus influenzae* (NTHi) were examined as a case-study. To this aim, a panel of lectins with diverse binding preferences was employed. The results revealed the availability of carbohydrate structures that could be targeted by endogenous lectins with the appropriate carbohydrate-binding specificity. The recognition of clinical isolates of different origin by innate immunity-associated lectins was also explored, yielding semi-quantitative data on binding avidity. In addition, to enable determination of binding affinity and kinetics for representative NTHi-lectin pairs, Quartz Crystal Microbalance (QCM) bacteria-based biosensors were designed, developed and used. Aiming towards ligand dissection among the various glycoconjugates present on the bacterial surface, the lipooligosaccharide (LOS), a major virulence factor, was isolated, quantified and structurally characterised by a combination of homo- and heteronuclear NMR techniques. Using designer microarrays, the LOS was found to be a target for some lectins, and in two selected cases the recognised epitopes were defined at atomic level using NMR-STD experiments. Finally, a pilot study of the influence of selected endogenous lectins on bacterial viability *in vitro* and on the potential for adhesion and invasion in cell culture was carried out.

Second, different extracellular vesicle populations were isolated and characterised by determination of their total protein cargo, particle concentration and morphology. The development of a widely applicable analytical tool enabling the glycoprofiling of distinct vesicular populations was a key objective. To this end, novel vesicle-based microarrays were set-up and validated by the use of selected known vesicle-binding molecules. Vesicle recognition by a panel of lectins of well-defined specificities was examined, revealing specific binding motifs. Based on lectin preferences for distinct populations, QCM chips were developed aimed at direct "fishing" and detection of vesicles in flow. In addition, the binding kinetics for representative vesicle-lectin pairs was investigated using vesicle-based QCM chips, here described for the first time.

Finally, a novel microarray application for the search of anti-adhesive compounds was developed. FimH, a carbohydrate-binding adhesin of *Escherichia coli* involved in urothelium colonisation and infection, was selected as target. A microarray set-up was designed for assessing FimH functionality on the surface of fixed and live bacteria at different growth phases. Furthermore, a competitive binding assay was established

using the same microarray platform and concanavalin A for set-up validation. Model multivalent ligands were tested as inhibitors of the binding to the arrays, and SPR and ITC analyses of the interaction served to confirm the microarray-derived results. The validated microarray set-up was finally used for evaluation of glycofullerenes as inhibitors of *E. coli* bacterial adhesion to array-printed ligands.

MAIN CONCLUSIONS

- Bacteria-based microarrays have proved to be efficient for lectin typing of bacterial glycosignatures, also providing semi-quantitative data on binding avidity. The results evidenced the availability of Man/Glc, sialic acid, and Gal residues on the surface of NTHi for lectin recognition. Consistent with this observation, the binding of lectins of the innate immune system with the appropriate carbohydrate-binding specificity, including Siglec-14 and several members of the galectin family, has been detected. In addition, the first experimental evidence for direct binding of SP-D to NTHi has been obtained. Comparison of the behaviour of selected galectins belonging to the three structural subgroups of this family unveiled striking differences in NTHi binding, which was found to be driven by carbohydrate recognition in the case of galectins 3 and 8, and by ionic interactions for galectin-4 and its C-terminal carbohydrate-recognition domain. LOS microarrays and NMR-based characterisation and mapping revealed lectin- and oligosaccharide-specific epitopes, detecting even slight divergences in the behaviour of lectins with similar nominal binding specificities, as the Gal-recognizing *Ricinus communis* and *Viscum album* agglutinins .
- Lectin screening of distinct EV populations immobilised into microarrays pointed to EV-specific glycosylation profiles that could be exploited for designing new isolation protocols based on lectin affinity chromatography.
- Pioneer QCM sensor chips incorporating entire bacteria or EV have been developed and used for label-free kinetic analysis of probe−lectin interactions directly at the surface, yielding quantitative data on binding affinity and kinetic parameters. In addition, a straight-forward QCM approach for in-flow selective capture of vesicles of interest has been set up, paving the way for non-invasive, high-throughput screening of aberrant glycosignatures transported by EV through biological fluids.

- Finally, a microarray set-up for assessment of the functionality of *E. coli* FimH adhesin has been successfully established and found to be efficient for discriminating between fimbriate (FimH-expressing) and afimbriate states in live bacteria. In addition, its potential for evaluation of dodecafullerenes as inhibitors of FimH binding has been demonstrated. Parallel microarray, SPR and ITC analyses using ConA as model lectinhave served to validate the microarray-derived results.

RESUMEN

PATRONES DE GLICOSILACIÓN Y RECONOCIMIENTO POR LECTINAS

INTRODUCCIÓN

La mayoría de los procesos de señalización biológica que tienen lugar en los organismos vivos involucran cuatro categorías de biomoléculas: ácidos nucleicos, proteínas, carbohidratos y/o lípidos. Los ácidos nucleicos y las proteínas, cuya secuencia y estructura fue objeto de estudio exhaustivo durante el siglo pasado, fueron identificados como biomoléculas responsables de la transmisión de información genética y como moléculas efectoras, respectivamente. No obstante, no parecía posible explicar integralmente procesos dinámicos a nivel intra- e intercelular, como por ejemplo la adhesión celular, la infección microbiana o la quimiotaxis de linfocitos, en un contexto de sólo dos dimensiones, *i.e.* ácidos nucleicos y proteínas. La existencia de un sistema encriptado adicional parecía esencial para cumplir estas funciones y los carbohidratos, ubicuos y diversos, se revelaron como biomoléculas idóneas para ejercer este papel.

La extraordinaria capacidad de biocodificación de la que disponen deriva de su versatilidad estructural. Para una determinada secuencia de azúcares existen multitud de posibles isoformas: basta con variar la configuración anomérica del enlace glicosídico (α- o β-), la forma de cada anillo, la posición de los enlaces, la presencia o no de sustituciones específicas (fosforilación, sulfatación, etc.) y/o la disposición lineal o ramificada de las unidades que componen dicha secuencia. Es evidente que las glicoformas constituyen un sistema muy flexible y además editable de manera dinámica, que goza al mismo tiempo de cierta independencia del genoma. Una colección de enzimas se encarga de su modificación, confiriendo a las células la capacidad de adaptarse constantemente a su entorno. A pesar de este proceso de adaptación, existen ciertos patrones de glicosilación, que constituyen una "huella dactilar", identificativa de su portador.

La capacidad de los carbohidratos para actuar como señales de reconocimiento hace necesaria la existencia de una precisa maquinaria que permita hacer este código operativo: las lectinas. Esta familia universal de proteínas une carbohidratos selectiva y reversiblemente. Aunque la afinidad de las lectinas por mono- y disacáridos es intrínsecamente baja ($K_D = 10^{-4}$ a 10^{-3} M), ésta se regula a diferentes niveles dando paso

a una notable selectividad hacia sus dianas. Por lo que se refiere a la lectina, la afinidad aumenta cuando ésta presenta sitios de unión extendidos, que ofrecen un área mayor de contacto, o múltiples sitios que operan simultánea y cooperativamente. Por otro lado, la presentación multivalente de ligandos, ya sea dentro de una misma molécula (por ejemplo, en oligosacáridos ramificados) o por agrupación de moléculas adyacentes, resulta en un aumento geométrico de la afinidad (efecto *cluster*).

En un contexto fisiológico, la matriz extracelular y el glicocálix, que se localiza en el exterior de la membrana de células eucariotas, patógenos o incluso derivados celulares (cuerpos apoptóticos, exosomas y demás vesículas) se caracterizan precisamente por presentar una alta densidad tridimensional de glicanos. En consecuencia, estas estructuras están enriquecidas en sitios de unión para lectinas endógenas, como son las lectinas del sistema inmune innato. Asimismo, los patrones de glicosilación específicos de cada una de ellas permiten su tipificación. En el caso de las bacterias, sus "glicofirmas" pueden contener carbohidratos que no se encuentran en el glicocálix de las células eucariotas, o también pueden imitar a los carbohidratos del huésped con el fin de evitar el reconocimiento por lectinas endógenas y así evadir la respuesta inmune. En las bacterias Gram-negativas, en particular, los lipopolisacáridos y los polisacáridos capsulares constituyen factores de virulencia importantes. Ciertas especies bacterianas (*e.g. Haemophilus, Neisseria,* etc.) presentan en la cara externa de su envoltura celular estructuras de lipooligosacárido (LOS), anclado a la membrana mediante el lípido A, y consistente en un núcleo interno y ramificaciones de unas pocas unidades sacarídicas en modo no repetitivo. La extensión de estas ramificaciones está sujeta a control de variación de fase, un proceso de activación genética selectiva que puede dar lugar a evasión inmune. Por tanto, el glico-fenotipado de patógenos junto a la identificación de los epítopos reconocidos por lectinas endógenas podría permitir el establecimiento de correlaciones de virulencia.

Para la identificación de los ligandos de lectinas endógenas se han utilizado técnicas de alto rendimiento (*high-throughput*), como por ejemplo plataformas de *microarrays* de carbohidratos. Éstos, sin embargo, presentan una serie de limitaciones, como es estar compuestos por librerías relativamente limitadas de oligosacáridos, generalmente de origen eucariota, cuya presentación en el *array* difiere sustancialmente con respecto a su entorno natural. Como consecuencia, estas plataformas no permiten evaluar la accesibilidad y densidad de cada oligosacárido en la superficie diana. Asimismo, ya que

sólo se examina la unión a ligandos individuales, tampoco puede detectarse una posible unión derivada del reconocimiento simultáneo de diferentes moléculas vecinas.

En el caso de las células eucariotas, los patrones de glicosilación se caracterizan por ser más estables y presentar escasos cambios fenotípicos individuales. No obstante, sí pueden verse alterados en casos de enfermedad, siendo el cáncer el ejemplo más prominente. Las células tumorales pueden presentar un perfil de glicosilación aberrante, tanto a nivel estructural como de expresión, en comparación con sus homólogos no transformados. La detección mediante procedimientos no invasivos de glicoproteínas asociadas a malignidad que pudieran ser de aplicación en diagnóstico temprano sigue siendo un reto difícil de conseguir, ya que estas glicoproteínas en muchos casos también se expresan en enfermedades benignas. Durante la última década, el descubrimiento de que las vesículas extracelulares pueden trasferir de manera horizontal (*i.e.* de una célula a otra) patrones moleculares, metabolitos o efectores, ha catalizado su uso como biomarcadores. Estas vesículas membranales son secretadas por casi todo tipo de células y se cree que comparten patrones de glicosilación con su célula de origen. Sin embargo, la glicofenotipificación de poblaciones individuales de vesículas extracelulares está dificultada en gran medida por la falta de metodología específica apropiada.

De todo lo anterior se desprende que la caracterización de los patrones de glicosilación de diferentes tipos de sondas, como bacterias o vesículas extracelulares, puede ayudar al establecimiento de correlaciones funcionales. Para ello es necesario el desarrollo de nuevas metodologías analíticas de alto rendimiento que integren el uso de lectinas como herramientas únicas de descodificación.

OBJETIVOS

El objetivo central de esta Tesis doctoral es el diseño y desarrollo de nuevas plataformas de alto rendimiento, destinadas al estudio de interacciones proteína-carbohidrato directamente en la superficie de bacterias y vesículas extracelulares, preservando el entorno y densidad natural de los epítopos. Los objetivos específicos son los siguientes:

- Desarrollo de *microarrays* de diseño y nuevas aplicaciones de la microgravimetría basada en cristales de cuarzo (Quartz Crystal Microgravimetry, QCM) para el análisis de la glicosilación de superficie de bacterias, y la exploración de su reconocimiento por lectinas del sistema inmune innato. Como bacteria modelo, se ha

utilizado *Haemophilus influenzae* no tipificable (NTHi), la cual es un ejemplo relevante de patógeno oportunista no encapsulado.

- Desarrollo de *microarrays* de vesículas extracelulares y chips de QCM como herramienta para el análisis de los patrones de glicosilación de exosomas y microvesículas de diferente origen y el estudio de los parámetros cinéticos de interacciones vesícula-lectina.

- Desarrollo de un método de *microarrays* que permita la evaluación de la funcionalidad de adhesinas bacterianas en bacterias vivas y el escrutinio de nanosistemas glicosilados como posibles agentes anti-adhesivos.

RESULTADOS Y APORTACIONES FUNDAMENTALES

En esta memoria se describe el diseño, validación y aplicación de nuevas estrategias para el glicofenotipado de bacterias y vesículas extracelulares, y para la evaluación de compuestos anti-adhesivos, empleando las tecnologías de *microarrays* y QCM.

En la sección 4.1, se describen los patrones de glicosilación obtenidos para distintas cepas de NTHi mediante ensayos de unión a *microarrays* de NTHi de un panel de lectinas con diferentes especificidades de unión a carbohidratos. Asimismo, se detallan los parámetros cinéticos obtenidos para la interacción de pares NTHi-lectina representativos, usando novedosos chips de QCM que incorporan la bacteria entera. La identificación del LOS de NTHi como ligando de ciertas lectinas ha dado paso a su análisis estructural por Resonancia Magnética Nuclear y a la elucidación a nivel atómico de los epítopos reconocidos mediante Diferencia de Trasferencia de Saturación (STD-NMR). Por otra parte, el estudio de la unión de lectinas del sistema inmune innato a diferentes aislados clínicos de NTHi ha demostrado un reconocimiento selectivo que depende tanto de la lectina como del aislado. Finalmente, se ha efectuado un estudio piloto de la influencia de lectinas endógenas seleccionadas sobre la viabilidad bacteriana *in vitro* y su capacidad de adhesión e invasión del epitelio pulmonar, no observándose efectos significativos.

En la sección 4.2, se describe el desarrollo de una aplicación de la tecnología de los *microarrays* que permite el glicotipado de poblaciones distintas de vesículas extracelulares. Las diferentes preparaciones de vesículas se han caracterizado mediante determinación del cargo total de proteína, medida de la concentración de las partículas y

observación morfológica por microscopía electrónica de trasmisión. Después de incorporar las vesículas en *microarrays,* sus patrones de glicosilación se han explorado mediante ensayos de unión del panel de lectinas de especificidad conocida. Al igual que se hiciera para las bacterias, se han desarrollado chips de QCM derivatizados con lectinas para la captura directa y detección selectiva de vesículas en flujo. Igualmente, se han realizado estudios cinéticos de la interacción de pares vesícula-lectina representativos, directamente sobre la superficie de las vesículas, empleando chips de QCM con vesículas enteras incorporadas.

Por último, en la sección 4.3, se describe una novedosa aplicación de los *microarrays* para la búsqueda de compuestos anti-adhesivos. En concreto, los *microarrays* desarrollados han permitido evaluar la funcionalidad en su entorno natural de FimH, una adhesina de *Escherichia coli* que reconoce carbohidratos e interviene en la infección de epitelios humanos, empleando para ello células de *E. coli* vivas. Se ha establecido un ensayo competitivo que permite la valoración de compuestos inhibidores de la adhesión, incluyendo glicoestructuras supramoleculares (fulerenos) o gliconanopartículas de silicio. La plataforma se ha validado empleando concanavalina A como lectina modelo, y los resultados obtenidos se han confirmado mediante resonancia de plasmón de superficie (SPR) y calorimetría de valoración isoterma (ITC).

CONCLUSIONES PRINCIPALES

- Se han desarrollado *microarrays* basados en la incorporación de bacterias completas que permiten su glicofenotipado, proporcionando datos semi-cuantitativos sobre la avidez de la unión de diferentes lectinas a su superficie. Los resultados han revelado la presencia de residuos de Man/Glc, ácido siálico y Gal en la superficie de NTHi, los cuales actúan como señales de reconocimiento para lectinas del sistema inmune innato, incluyendo Siglec-14 y diferentes galectinas. Además, se ha obtenido la primera evidencia experimental del reconocimiento directo de NTHi por SP-D. El estudio comparativo del comportamiento de galectinas seleccionadas, pertenecientes a los tres subgrupos estructurales de esta familia de lectinas, ha revelado claras divergencias. En el caso de las galectinas 3 y 8, la unión a NTHi está mediada por el reconocimiento de carbohidratos. Sin embargo, la unión de la galectina-4 depende fundamentalmente de interacciones iónicas. Utilizando microarrays de LOS y diferentes técnicas de RMN, se han

identificado los epítopos reconocidos por determinadas lectinas, detectándose interesantes diferencias en el comportamiento de lectinas homólogas, como es el caso de las aglutininas específicas de Gal procedentes de *Ricinus communis* y *Viscum album*.

- El estudio del reconocimiento de distintas subpoblaciones de vesículas extracelulares por un panel de lectinas con distintas especificidades de unión ha demostrado la existencia de perfiles de glicosilación característicos para cada subpoblación, los cuales podrían ser aprovechados para diseñar nuevos procedimientos para su fraccionamiento.
- Se han desarrollado nuevos chips de QCM que incorporan bacterias o vesículas extracelulares enteras, y permiten el análisis de las interacciones sonda-lectina, directamente en su superficie, proporcionando datos cuantitativos sobre la afinidad de unión y parámetros cinéticos. Por otro lado, se ha establecido una nueva y sencilla aproximación que permite la captura selectiva en flujo continuo de vesículas de interés, abriendo así el camino para la detección y escrutinio no invasivo y de alto rendimiento de glicofirmas aberrantes transportadas por vesículas extracelulares en fluidos biológicos.
- Por último, se han desarrollado nuevos *microarrays* de diseño para la evaluación de la funcionalidad de la adhesina FimH de *E. coli* (específica de mannosa), los cuales permiten discriminar entre bacterias fimbriadas y no fimbriadas. Adicionalmente, se ha demostrado su utilidad para la evaluación de glicofulerenos como inhibidores de la adhesión bacteriana. La validación de la plataforma utilizando concanavalina A como lectina modelo ha incluido el análisis de las interacciones lectina-inhibidor mediante resonancia de plasmón de superficie y calorimetría de valoración isoterma, confirmando los resultados obtenidos.

LIST OF TABLES

Chapter 1: Introduction

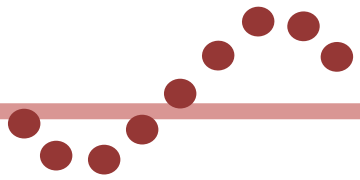
Chapter 3: Materials & Methods

Chapter 4: Results

LIST OF ABBREVIATIONS

AFET	Asialofetuin
ATCC	*American Type Culture Collection*
BSA	Bovine serum albumin
CFU	Colony forming unit
ConA	Concanavalin A
CRD	*Carbohydrate Recognition Domain*
DC-SIGN	*Dendritic Cell-Specific Intercellular adhesion molecule-3 Grabbing Non integrin*
DMSO	Dimethyl sulfoxide
DNase	Deoxyribonuclease
DOC	Sodium deoxycholate
EBSS	*Earle's Balanced Salt Solution*
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EXO	Exosomes
FBS	Fetal bovine serum
FET	Fetuin
Fc	Immunoglobulin fragment C
Fg	Fibrinogen
FITC	Fluorescein
Fuc	Fucose
GAG	Glycosaminoglycan
Gal	Galactose
GalNAc	N-acetyl-galactosamine
Glc	Glucose
GlcNAc	N-acetyl-glucosamine
GM3	Mono-sialo-di-hexosyl-ganglioside
GNP	Glyconanoparticle
GSL	Glycosphingolipid
HBS	10 mM HEPES buffer, pH 7.4, 0.15 M NaCl
HBSC	HBS containing 10 mM $CaCl_2$
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hex	Hexose
Hep	L-glycero-D-manno-Heptose
Hz	Herz, frequency unit
hGal	Human galectin
IC_{50}	Half maximal inhibitory concentration
IgG	Immunoglobulin G
ITC	*Isothermal Titration Calorimetry*
k_a	Association rate constant
k_d	Dissociation rate constant
K_D	Dissociation constant
K_{eq}	Binding (equilibrium) constant
KdO	3-deoxy-D-manno-oct-2-ulosonic acid
Lac	Lactose, Galβ(1,4)Glc
LB2YT	Enriched Lysogeny Broth medium, containing double yeast extract and tryptone
LOS	Lipooligosaccharide

LPS	Lipopolysaccharide
Man	Mannose
Manno-Hep	L-glycero-D-manno-Heptose
MOI	Multiplicity of infection
MVs	Microvesicles
NC	Nitrocellulose
Neu5Ac	N-acetyl-neuraminic acid
NHS	N-hydroxysuccinimide
NMR	*Nuclear Magnetic Resonance*
NTHi	Non-typeable *Haemophilus influenzae*
OD	Optical density
OS	Oligosaccharide
PAGE	*Polyacrylamide Gel Electrophoresis*
PBS	5 mM sodium phosphate buffer, pH 7.2, 0.2 M NaCl
PBS-d	Deuterated PBS
PBS_{β}	PBS containing 4 mM *β*-mercaptoethanol
PBS_{DTT}	PBS containing 2 mM dithiothreitol
PDB	*Protein Data Bank*
PPEtn	Pyrophosphorylethanolamine
PRR	*Pattern-Recognition Receptor*
rGal	Rat galectin
rfu	Relative fluorescence unit
RI	Refractive index
RNase	Ribonuclease
RPMI-1640	*Roswell Park Memorial Institute* medium
QY	Quantum yield
QCM	*Quartz Crystal Microgravimetry*
sBHI	Supplemented Brain Heart Infusion medium
SDS	Sodium dodecylsulphate
SNP	Silica nanoparticle
SP-D	Human surfactant pulmonary-associated protein D
SPR	*Surface Plasmon Resonance*
SPW	Surface plasmon wave
STD	*Saturation Transfer Difference*
TBS	10 mM Tris, pH 7.8, 0.15 M NaCl
TEM	Trasmission Electron Microscopy
THP-1	Human acute monocytic lymphoma cell line
TRPS	Tunable Resistive Pulse Sensing
UTVG	Unsubstituted terminal vicinal glycol
UV-vis	Ultraviolet-visible
wt	Wild type
α-Me-Man	*a*-methyl-mannoside
*β*ME	*β*-mercaptoethanol
Δf	Frequency shift
ΔG°	Gibbs free energy
ΔH	Enthalpy change of the binding event
ΔS	Entropy change of the binding event



INTRODUCTION

1.1 The Codes of Life

During the second half of the 20th century, the leading lines of research focused on unravelling the chemical basis of heredity. Proteins were the first cell components to be considered the genetic information carrier, because of being both abundant and complex, made up of twenty main building blocks (amino acids). Up until O. Avery *et al.* discovered that DNA is the essential molecule able to induce bacterial transformation (Avery *et al.*, 1944), it had been considered a rather monotonous substance, not intricate enough to convey a huge amount of information, as it is only constituted by four alphabet characters: adenine, thymine, guanine and cytosine. A few years later, Watson and Crick observed that complementarity of the hydrogen-bonding patterns of the four bases is critical not only for DNA stability but also for its replicability (Watson and Crick, 1953). Both properties serve the necessity of rigidity and optimal fidelity in the inherited genetic information.

From one aspect, adding letters to this alphabet would bring about more sequence combinations of nucleotides and thus increase the informational content of oligomers. Nevertheless, summing up manifold chemical, thermodynamic and computational studies, four still seems to be the optimal size for the genetic alphabet, "as a compromise between stability and evolvability, fidelity and enzymatic/catalytic efficiency as well as information density and error resistance" (Szathmary, 2003).

The genetic code, relating two sequence types (i.e. nucleic acids into proteins) in a linear manner via the transcription and translation processes, triggers the conversion of triplet coding to peptides arranged in bioactive shapes (Barbieri, 2008). As opposed to DNA's rigid structure, the higher degree of limberness of peptides can endow proteins with the capacity to adopt distinct functional conformations or even to behave as conformational switches. Moreover, proteins are built up by twenty main different "letters" and, as a consequence, they constitute an information carrier of higher intricacy.

Furthermore, the ubiquitous presence of co- and post-translational modifications, such as phosphorylation or glycosylation, suggests that a protein's activity profile is determined at different levels, beyond its sequence. Of course, the covalent attachment of glycans can directly influence the macromolecule´s intrinsic properties, such as solubility, molecular mass or electric charge. It may also have a significant impact on the minimisation of misfolding by offering assistance, thus playing an important role in the protein quality control, a fundamental cellular activity (Roth *et al.*, 2010).

What is more, glycosylation produces an abundant, highly regulated inventory of conjugates containing protein or lipidic cores, which can be localised intracellularly, on the cell membrane or in the extracellular matrix (see below). Using glycan-binding tools and high-throughput multiplex analysis techniques, glycophenotyping of cells and tissues can nowadays be conveniently accomplished. Glycomes have been proved to be highly cell specific, hinting at their utility as unique cellular descriptors (Fujitani *et al.*, 2013) or molecular fingerprints, the glycosignatures.

To realize all the potential of carbohydrates for variability, a refined toolbox of enzymes is dedicated to glycan biosynthesis and modification, and biomolecule glycosylation, following an actively regulated pathway. Factors including the *in situ* concentration of substrates as well as the presence and localisation of distinct glycosidases and glycosyltransferases confer to the process a decidedly dynamic character and add up to determine the glycophenotype. An intriguing sensitivity to epigenetic influences as well as infections characterizes the latter and attests that the "Sugar Code" is not directly encoded by the genome, hereby conferring to living cells adaptability and capacity of response to environmental stimuli.

1.2 The Sugar Code: the "sweet" side of molecular recognition

For a long time, carbohydrates, the most abundant type of biomolecules in nature, were almost entirely side-lined with respect to biological information transfer, and had been assigned the role of building blocks of protective cell-wall constituents (e.g. cellulose and chitin) or metabolism fuel (e.g. glycogen). However, this aspect was questioned occasionally over the years and, in the '70s, the landscape started to change, when the oligosaccharides of mammalian tissues and serum glycoproteins were assigned functional roles (Laurence *et al.*, 1972).

At the same time, attempts to comprehensively understand intra- and intercellular recognition processes, such as lymphocyte routing and cell adhesion, seemed unviable in the framework of only two dimensions (nucleic acids and proteins), Fig. 1.1. The consideration of another encrypted system was therefore essential to bridge this gap, and sugars, with their unique versatility and dense bio-coding capacity, had already started to reveal themselves as very "shapely" molecules (Rees, 1972).

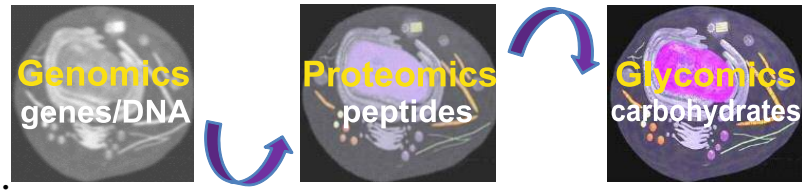


Fig. 1.1 From genomics to glycomics passing through proteomics: each step brings us closer to a clearer view of molecular and cellular physiology.

Essentially, five are the structural characteristics of sugars that confer them an outstanding aptitude for intricate coding: their potential for variability in (i) the anomeric configuration of the glycosidic linkage (α or β), ii) the position of this bond (1 → 1, 2, 3, 4, 6 for hexopyranoses), iii) the ring size (five-membered, furanose, or six-membered, pyranose rings), iv) site-specific substitutions (phosphorylation, acetylation or sulphation), and v) the linear or branched arrangement of oligomers (Gabius *et al.*, 2004).

For a trisaccharide composed of 3 random sugars out of more than 20 commonly found in mammals (i.e. glucose, mannose, galactose, fucose, *N*-acetyl-glucosamine, etc.), the number of occurring trimers by sequence permutation alone reaches 20^3 (8000). However, as more factors should be considered in a trisaccharidic structure, the number of possible occurring molecules should be calculated as [*Permutations* × *anomerics* × *ring sizes* × *linkages*] what results in more than 6 million structures. Hence, oligomers of more than 8 constituent parts, which also have the potential of generating numerous branched isomers, would sum up to astronomical numbers, superior to the Avogadro's number (Laine, 2008).

Furthermore, the described carbohydrate structures, either linear or branched, encompass only the first and second dimensions of the Sugar Code. The conformational behaviour of oligosaccharides, also determined by stereo-electronic effects and steric interactions across the glycosidic bond (Solís *et al.*, 2015), constitutes the third dimension. The twist angles around the glycosidic linkage can vary, so that a carbohydrate structure of well-defined sequence can adopt several conformations in solution in which the relative orientation of the monomers is different. Questions on low-energy conformations of oligosaccharides, flexibility and dynamics can be answered experimentally, using e.g. high-field nuclear magnetic resonance, while

several *in silico* tools and methods have been developed in order to assist in the interpretation of experimental results (Lütteke and Frank, 2009).

In an evolutionary frame, how could organisms take advantage of such a convoluted coding system? In order to answer this question, we have to consider that without the constraint of having to achieve optimal fidelity, a bio-code is expected to reach a high information density. Conformational flexibility plus the potential for rapid and reversible structural modulations by e.g. phosphorylation/sulphation are another two strategic elements of a powerful coding tool (Gabius, 2000). All the above requisites are met in the Sugar Code, together with the gain of relative genome independence.

More to the point, the space on the cell surface is fairly limited and has to harbour a multitude of signals, making high-density coding necessary. Indeed, most cells have a fuzz-like carbohydrate coat on the external surface of their plasma membranes named glycocalyx (Bennett, 1963), meaning "sugar coat" (from greek *γλυκύς* = sweet and *κάλυξ* = husk). Tissue cells are also surrounded by a dynamic and complex-nature resilient meshwork. The extracellular matrix provides mechanical and structural support, but also the spatial context for signalling events to take place, with direct implications in developmental patterning, maintenance of stem cell niches and cancer progression (Hay, 1991; Kim *et al.*, 2011). Sugars constitute an important part of both the calyx and the extracellular matrix and are linked to distinct aglycon parts to form a variety of glycoconjugates. Glycolipids, glycoproteins and proteoglycans are the most representative families (please see Fig. 1.2). In detail:

- *Glycolipids* are ubiquitous lipid–glycan conjugates usually found in the plasmatic membrane clustered in microdomains, the "lipid rafts", with their glycosidic part protruding towards the extracellular space, freely accessible for intermolecular interactions. They contain one or more monosaccharides bound to a hydrophobic membrane-anchoring compound such as an acylglycerol, a sphingoid, a ceramide (*N*-acyl-sphingoid) or a prenylphosphate. Most mammalian glycolipids belong to the class of glycosphingolipids (GSLs) and contain a lipid backbone composed of sphingosine and a fatty acid. Glucosylceramide is the glycolipid core most frequently found in mammals, and is the starting point for the biosynthesis of GSLs of higher glycan complexity by Golgi-localised glycosyltransferases. GSLs are divided into two categories – neutral or

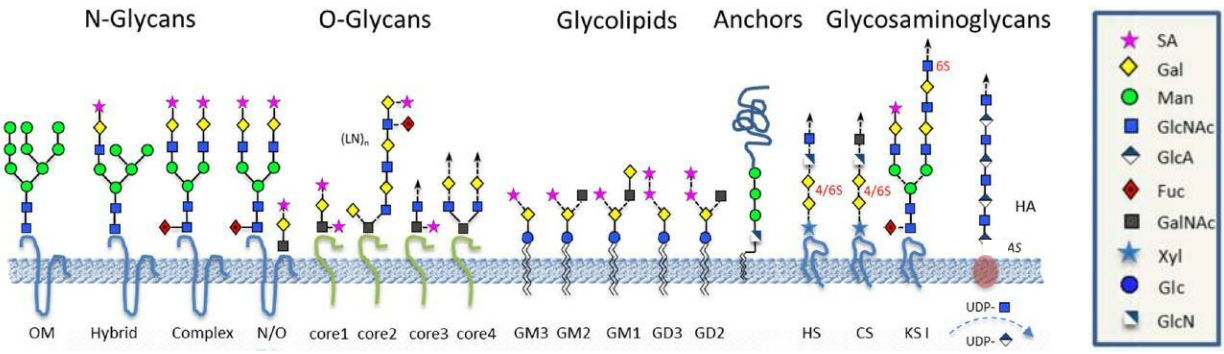


Fig. 1.2 Schematic depiction of the various classes of mammalian glycans. From left to right: OM, oligomannose; GM and GD, mono- and disialylated glycosphingolipids; HS, heparan sulphate; CS, chondroitin sulphate; KS I, keratan sulphate; HA, hyaluronan. Monosaccharide abbreviations are as follows: SA, sialic acid; Gal, galactose; Man, mannose; GlcNAc, *N*-acetyl-glucosamine; GlcA, glucuronic acid; Fuc, fucose; GalNAc, *N*-acetyl-galactosamine; Xyl, xylose; Glc, glucose; GlcN, glucosamine. Adapted from (Dalziel *et al.*, 2014).

acidic – depending on whether they carry or not a negative charge due to sulphation or sialylation (Kopitz, 2009).

- *Glycoproteins* contain oligosaccharide chains, which can be either *N*-linked or *O*-linked. *N*-glycan biosynthesis in mammals starts with the transfer of the preassembled lipid-linked oligosaccharide $Glc_3Man_9GlcNAc_2$ to an asparagine residue. Then, the oligosaccharide undergoes a step-wise size reduction, in some cases reaching a pentasaccharide stage ($Man_3GlcNAc_2$) followed by a subsequent neo-synthesis, thereby contributing to glycan remodelling and opening the way to the complexity of the cellular *N*-glycome. Besides mannose, *N*-acetyl-glucosamine and galactose, *N*-glycan chains may also contain L-fucose and *N*-acetyl-neuraminic (sialic) acid, or get substituted by sulphate or phosphate groups (Zuber and Roth, 2009). In addition, the hydroxyl groups of serine/threonine also serve as potential targets for glycan attachment, with *O*-glycan sites often appearing as clusters. *O*-glycans are synthesised by transfer of one monosaccharide (GalNAc, GlcNAc, Man, Fuc or Glc) to the site and sequential addition of others, forming a core unit that can be later extended (Patsos and Corfield, 2009).
- *Proteoglycans* (PGs) make up a family of complex macromolecules consisting of a protein core to which different high-molecular-weight polyanionic glycosaminoglycan (GAG) chains are linked covalently. GAGs are linear polymers of 50-150 repeating disaccharides built up by a hexosamine and a uronic acid, usually D-Glucuronic (D-GlcA) or L-Iduronic acid (L-IdoA), or galactose (Esko *et al.*, 2009). GAG chains are attached to the protein core via a carbohydrate–protein linkage region which is either a tetrasaccharide (GlcA-Gal-Gal-Xyl) glycosidically linked to a serine residue or has a

glycoprotein-like structure. Sulphate groups are usually present on carbon atoms 4′ or 6′ of the hexosamine entity, but could also be present on C2 of L-IdoA. GAG saccharide chains usually correspond to a molecular weight greater than 20 kDa (Buddecke, 2009). These simple definitions encompass an exceptionally large range of structures involving different core proteins (molecular weight ranges 20-250 kDa) as well as diverse classes or numbers and lengths of individual GAG chains (chondroitin/heparan/keratan sulphate, etc.) (Wight *et al.*, 1991), depending on selective enzyme expression and localisation (Esko *et al.*, 2009).

In addition to offering chemical resistance and backbone support, this extended presence of saccharides on the cellular surface plays an important functional role, frequently mediating cell recognition and adhesion processes, such as tissue organization, fertilization (Töpfer-Petersen *et al.*, 1998), growth regulation (Iozzo, 2005; Villalobo and Gabius, 1998), immune response (Rabinovich and Toscano, 2009), inflammation (Jack *et al.*, 2001; Schwartz-Albiez, 2009) or cancer (Hakomori, 1985; Ohtsubo and Marth, 2006). This is why specific alterations of the glycolipid/glycoprotein glycosylation patterns or up-regulation of their synthesis have recently been associated to pathogenesis, suggesting certain structures as disease markers or potential targets for therapeutic intervention (Dalziel *et al.*, 2014; Dube and Bertozzi, 2005; Peracaula *et al.*, 2008). Besides, carbohydrates also constitute markers for antibodies (Calarese *et al.*, 2005), hormones (Villalobo and Gabius, 1998), plant (Jiménez *et al.*, 2008; Kabat *et al.*, 1947; Olsnes *et al.*, 1978) or bacterial toxins (Haan and Hirst, 2004), and viral (Childs *et al.*, 2009) or bacterial receptors (adhesins) (Karlsson, 1995; Krogfelt *et al.*, 1990).

In particular, aiming at actively inducing their entry into target cells for replication and/or immune disguise, both Gram-positive and -negative bacteria have evolved sophisticated molecular strategies and machinery, frequently involving specific recognition of extracellular matrix components (Kline *et al.*, 2009). Direct engagement of these elements by bacterial adhesins (Remaut and Waksman, 2004) or translocation of bacterial effector molecules into the target-cell cytosol was shown to promote rearrangements of the plasma membrane architecture and to facilitate pathogen engulfment (Mulvey, 2002; Pizarro-Cerdá and Cossart, 2006). Adhesins are anchored on the bacterial cell-wall, as the SdrG protein of *Staphylococcus epidermidis*, which binds to the Bβ chain of human fibrinogen (Ponnuraj *et al.*, 2003), or constitute components of supramolecular fibres

called fimbriae or pili, as FimH of uropathogenic *Escherichia coli*, UPEC (Bouckaert *et al.*, 2005; Choudhury *et al.*, 1999).

Of the various UPEC adhesins, type-1 pili FimH adhesins are by far the most prevalent (Bahrani-Mougeot *et al.*, 2002), being expressed by a wide range of pathogenic, commensal and laboratory strains (Fukiya *et al.*, 2004). FimH is exposed at the their tip, its carbohydrate-binding pocket being localised at the distal end of the protein N-terminal domain (Bouckaert *et al.*, 2005). It selectively binds a variety of mannose-presenting glycoproteins expressed by diverse host cell types (Ofek *et al.*, 1977), including bladder epithelial cells (Fig. 1.3). Uroplakin Ia, bearing oligomannose structures and abundantly present on differentiated urothelial cells, has been identified as target (Xie *et al.*, 2006; Zhou *et al.*, 2001). Although FimH binds to uroplakin Ia with moderate strength, the interaction of multiple fimbriae of an individual bacterium and the array of polymerised uroplakin on the cell surface lead to high avidity and ensure stable bacterial adhesion (Zhou *et al.*, 2001). The increasing incidence of bacterial resistance to antibiotics has stimulated the interest in the synthesis of anti-adhesive, sugar-bearing multivalent scaffolds to be used as alternative or complementary therapy, willing to overcome these intrinsic avidity issues (Reichardt *et al.*, 2013).

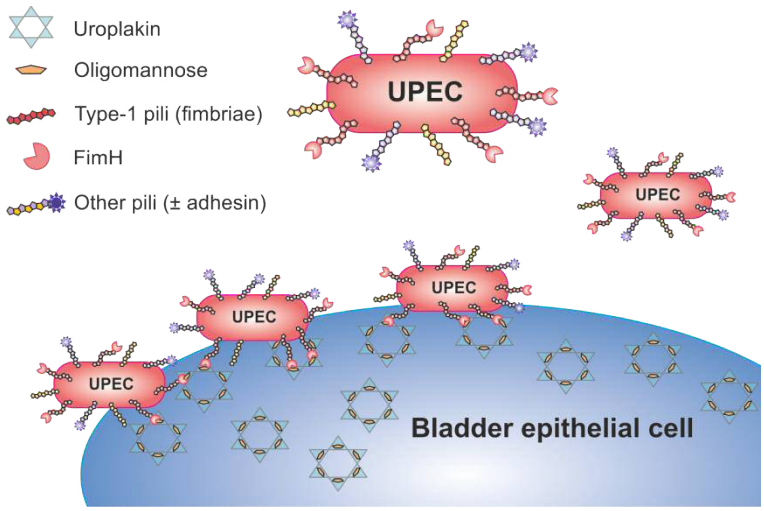


Fig. 1.3 Uropathogenic *E.coli* (UPEC) bears fibrous pili, frequently exposing adhesins at their tip. FimH recognises high-mannose structures on uroplakin Ia, a bladder epithelium glycoprotein. Bacteria strongly adhere on target cells by establishing multivalent interactions.

Apart from eukaryotic cells, pathogens are also coated with a variety of carbohydrate-rich structures that usually serve to typify strains (Peterson and Quie, 1981) and are specifically recognised by host receptors for triggering defence responses (Davicino *et*

al., 2011; Kumar *et al.*, 2011; Sahly *et al.*, 2008). In the case of bacteria (Fig. 1.4), lipopolysaccharides (LPSs) and capsular polysaccharides are important virulence factors often including unusual carbohydrate monomers with respect to mammals (Herget *et al.*, 2008; Schirm *et al.*, 2005). LPSs constitute one of the major outer-membrane components of Gram-negative bacteria, anchored to the outer leaflet of the bacterial cell envelope.

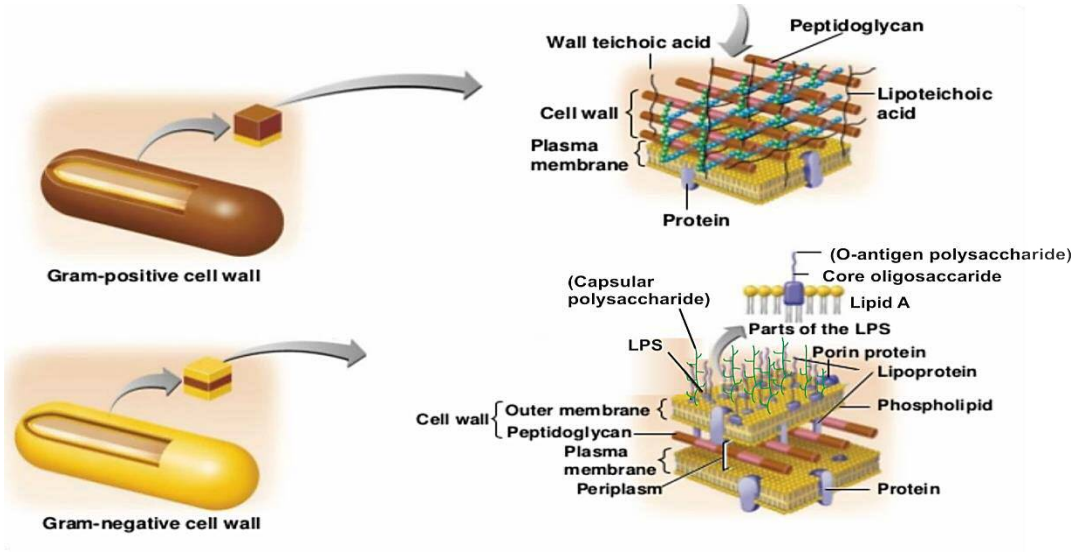


Fig. 1.4 Structural features of the bacterial cell walls, including peptidoglycan found in both Gram-positive and -negative bacteria, LPS and capsular polysaccharide. O-antigen and capsular polysaccharides are named in parenthesis, not being ubiquitous in Gram-negative species. Adapted from Midlands Technical College online courses, (©Pearson Education).

They are composed of a highly conserved lipid A molecule – serving as hydrophobic anchor – and a variable polysaccharide, the O-antigen/chain, built up by repeating saccharide units on a small inner core oligosaccharide (Freudenberg and Galanos, 1990). The core saccharide is connected to the lipid A moiety through one or more acidic sugars identified as 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo), and might be further subjected to assorted non-stoichiometric substitutions, such as addition of phosphate, phosphoethanolamine (PEtn), or pyrophosphoethanolamine (PPEtn) moieties (Gibson *et al.*, 1993). Some bacteria synthetize lipooligosaccharide (LOS) molecules characterised by the presence of an extended branched core region prominently composed of hexoses (Hex) and heptoses (Hep), but lacking repetitive O-chains (Phillips *et al.*, 1992).

Extensive structural studies of the LOS of *Haemophilus spp.* have led to the identification of the conserved inner-core moiety l-*α*-D-Hep*p*-(1→2)-[PEtn→6]-l-*α*-D-Hep*p*-(1→3)-[*β*-D-Glc*p*-(1→4)]-l-*α*-DHep*p*, linked to one Kdo 4-phosphate unit (Schweda *et al.*, 2007). Each backbone Hep can be a point for the addition of a Hex or further chain extensions,

which constitute the outer core. It is likely that this heterogeneity, including structural intermediates as well as products of alternative biosynthetic pathways, enables pathogenic bacteria to adapt to changing and potentially hostile environments, through rapid phase variation of their LPS structures (Morey *et al.*, 2013; Raetz *et al.*, 2007).

Glycosylation alterations are frequent in bacteria as well, and directly affect the virulence capacity of microorganisms either by promoting resistance to antimicrobial peptides and getting involved in self-aggregation and biofilm formation (Morey *et al.*, 2013), or by interfering with the ability of the host receptors to recognize LPSs as non-self elements (Raetz *et al.*, 2007).

Recently, bacteria have also been found to carry heavily glycosylated proteins, which may impede host inflammatory and immune responses (Schmidt *et al.*, 2003) or play specific roles in adhesion and colonisation (Mell *et al.*, 2016). Major O- and N-glycosylation pathways involve the synthesis of a lipid-linked oligosaccharide and its transfer *en bloc* to the protein acceptor by an oligosaccharyltransferase (Iwashkiw *et al.*, 2013; Nothaft and Szymanski, 2010). In an alternative pathway, glycosyltransferases sequentially transfer monosaccharides to the protein acceptor from their nucleotide-activated forms. This pathway, found to be operative for N-glycosylation only in a few γ-proteobacteria species (Naegeli *et al.*, 2014a; Naegeli *et al.*, 2014b), was first described in nontypeable *H, influenzae* (NTHi): NTHi HMWC is an inverting N-glycosyltransferase that adds Glc and Gal monosaccharides at Asn residues and is also able to form di-hexoses at the sites with Glc modification (Grass *et al.*, 2010; Gross *et al.*, 2008). Whereas the HMW1C-like glycosyltransferase from *Actinobacillus pleuropneumoniae*, the best studied member of this family at biochemical level, is able to modify different polypeptide structures *in vitro* (Schwarz and Aebi, 2011), the only endogenous substrate proteins described so far for NTHi HMWC are the high-molecular adhesins HMW1A and HMW2A (Gawthorne *et al.*, 2014; Naegeli *et al.*, 2014a).

In summary, carbohydrates have exceptional properties for storing biological information and potentially establishing intermolecular interactions, thereby contributing to multiple biological events. Yet, their in-principle advantageous complexity poses the new requirement of a multifaceted machinery specialised in making this coding operative. In fact, the mechanism that permits organisms to exploit all three dimensions of the

carbohydrates' potential is their recognition by protein receptors. The proteins capable to specifically recognise sugars and hence decipher the glycocode are the lectins.

1.3 Lectins, the translators of the Sugar Code

The first documented observation of lectin activity was made in 1861, when S.W. Mitchell reported blood "clotting" by the rattlesnake venom (Mitchell, 1861). Experiments were repeated with washed erythrocytes, in order to exclude interaction with other plasma components. The development of the classical haemagglutination assay addressed this necessity. This assay has proved to be instrumental to detect the cell cross-linking capacity of proteins from plant extracts. In fact, the first study on such an activity dates back to 1888, when Hermman Stillmark described the haemagglutination phenomenon induced by a protein factor present in *Ricinus communis* seeds, denominated ricin (Stillmark, 1888). Stillmark's study also revealed a certain degree of selectivity in agglutination of red blood-cells from different animal species. Meanwhile, the further discovery that plant extracts are rich sources of agglutinins made possible the first purification of such a protein (concanavalin A, ConA) by crystallisation (Sumner, 1919) and the demonstration of its specific interaction with carbohydrate groups by the precipitation of starch, glycogen and mucin (Sumner and Howell, 1936).

A few years later, the ability of ricin and other haemagglutinins to distinguish among ABH blood groups was evidenced and, since this behaviour resembled that of serum antibodies, these proteins were called antibody-like substances, due to their exquisite binding specificity. However, as Boyd explained in his classical paper in 1954 (Boyd and Shapleigh, 1954), "it would appear to be a matter of semantics as to whether a substance not produced in response to an antigen should be called an antibody, even though it is a protein and combines specifically with a certain antigen only". Although sugar-binding proteins are known to be up- or down-regulated in a variety of inflammatory processes, their biosynthesis does not require a previous immune stimulation. Moreover, they don´t present any structural resemblance to antibodies – with the exception of animal I-type lectins (Crocker *et al.*, 2007) – and they cannot chemically modify the carbohydrates they bind to, making their distinction from carbohydrate-active enzymes clear. Thus, Boyd introduced a new term and called those proteins *lectins* (from latin *lectus*, the past principle of *legere* meaning to pick, choose or select), due to their extraordinary capacity to select their ligand and specifically bind to it (Boyd, 1954). Still, in order to reach the present

version of the term *lectin*, its definition underwent several refinements. The experimental focus put initially on agglutination required at least bivalency for the cross-linking of two cell surfaces, thus excluding monovalent proteins. Subsequently, three criteria have been established that must be simultaneously met by a protein in order to qualify as a member of the lectin family: (i) confirmed carbohydrate-binding activity, (ii) clear distinction from immunoglobulins, and (iii) clear distinction from enzymes tailoring free saccharides and from sensor/carrier proteins for the same molecules.

Another significant discussion started with the realisation that a name for a protein does not often cover all its structural facets, as many proteins are known to be composed of several domains (modules). Such a modular design underlies operativity in multiple aspects and enables that a lectin domain can be linked to other functional modules, even including catalytic ones. Class I plant chitinases, for instance, are built by a chitin-binding hevein domain and a catalytic domain, separated by a hinge region (Collinge *et al.*, 1993). The protein modules´ association is physiologically relevant in terms of inter-domain cooperation, as found in numerous cases of endogenous receptors. As significant example relevant to the present work, chimera-type galectin-3 or pulmonary surfactant protein-D (SP-D) (Crouch, 1998) are instructive cases of proteins with one carbohydrate-recognition domain (CRD) linked to other types of bioactive regions (Fig. 1.5). These two proteins possess collagen-like tails available for auto-association and multimer formation, leading to suited CRD spatial vicinity and high-affinity ligand binding, pertinent to their physiological role (Brewer *et al.*, 2002; Dam and Brewer, 2008).

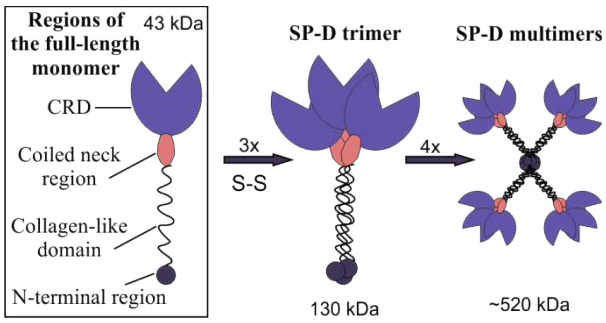
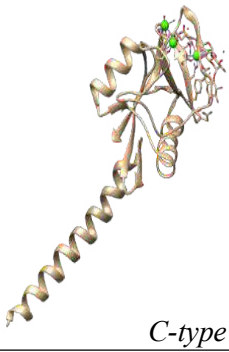
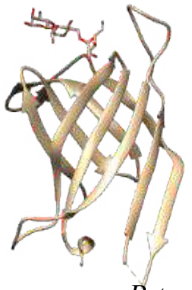


Fig. 1.5 Structure of SP-D monomers, trimers and multimers.

Following the concept of the Sugar Code, the frequency with which domains or protein folds are adapted to accommodate glycan epitopes as ligands should be fairly high in order to match the evident glycome complexity (Solís *et al.*, 2015). The past decades have witnessed the emergence of an insight into the structural and functional diversity of animal lectins. The establishment of several techniques to detect them and characterise their

Table 1.1 Principal families of animal lectins

Family	Overall fold	Specificity	Examples	Localisation	Functions
C-type	 *C-type*	Diverse	Collectins (MBL: **1AFB**, SP-D: **2GGU***) Asialoglycoprotein receptor: **1DV8** P-selectin: **1G1Q**	Extracellularly Cell membrane	Recognition of foreign/aberrant glycosignatures on cells (endocytosis or opsonisation) and of apoptotic/necrotic cells (glycans & peptide motifs).
I-type	 *Ig-like*	Sialic acids	Siglecs Sialoadhesin: **1QFO***) Neural Cell Adhesion Molecule (NCAM): **1EPF**	Cell membrane	Cell adhesion and growth, cell-cell interactions, recognition of microbial and host-like epitopes, induction/suppression of effector release and inflammatory cascades.
P-type	 *P-type*	Man-6-PO_4	OS-9 MRH domain: **3AIH*** Mannose-6P-receptor: **1M6P**	Intracellularly (endosomes, Golgi)	Intracellular routing of glycoproteins and vesicles, apical delivery.
S-type (galectins)	 *β-sandwich*	β-Gal	Animal, fungal, human galectins (hGal-1: **1GZW*** hGal-3: **1A3K** hGal-8: **3AP4**)	Extracellularly Intracellularly (cytoplasm, nucleus)	Extracellular matrix network assembly, cell migration and routing, cell-cell interactions, ligand-receptor bridging and endocytosis via lattice formation.
Pentraxins	 *β-sandwich*	Diverse	Serum amyloid P-component: **1GYK*** C-reactive protein: **1GNH**	Extracellularly	First-line of defence, recognition of foreign/aberrant glycosignatures (endocytosis or opsonisation).

Adapted from (Solís *et al.*, 2015).

* Protein Data Bank (PDB) files selected for graphical representation.

specificity and interactions with saccharides has paved the way for the dynamic development of the field of functional glycomics. Up to the present, 14 different folds have been identified in human/animal lectins (Gupta, 2012), many of which resemble folds of plant (Solís *et al.*, 2015), bacterial or viral lectin modules (Holgersson *et al.*, 2009), attesting their ubiquitous presence. For example, the β-sandwich fold was first observed in the crystal structure of ConA, the first to be solved for a leguminous lectin (Hardman and Ainsworth, 1972), although it is very common among animal lectins. This fold also proved out to be versatile enough to let several lectin groups arise, including galectins, pentraxins, calnexins, etc. Taken together the fairly large number of folds with the high level of intra-family diversification clearly argues in favour of the extraordinary potential of sugar receptors to be physiologically significant in a multitude of manners. Common folds and structure-based classification of the main animal lectin families are shown in Table 1.1, along with a brief description of their biological functions. The presence of at least one noncatalytic, carbohydrate-recognition domain reversibly binding to a specific carbohydrate is the denominator in the whole lectin family in all kingdoms, regardless of the presence or not of other cooperating modules (Peumans and Van Damme, 1995). What is more, all modes of spatial ligand accommodation that can be envisioned have co-evolved, so that the translation of the sugar-encrypted message into a physiological response is resolutely connected to the recognition capacity of the lectins. This selectivity is defined by different degrees of sugar engagement in the corresponding sites, following a theme of contact complementarity: starting from flat surfaces and shallow grooves it progresses to a firm grip of higher degrees of H-bonding to reach high affinity for di- or oligosaccharides (Gabius *et al.*, 2011).

1.4 Atomic features and thermodynamics of lectin-carbohydrate interactions

As described before, when directly compared to oligonucleotides and peptides in terms of coding capacity, oligosaccharides constitute an intricate lexicon characterised by the exceptional structural diversity of sugars. A wide panel of ligands (sugar-code words), both charged and neutral, is recognised by the endogenous lectins. A detailed insight into the atomic features of protein-sugar interactions is fundamental to understand how this complexity and high coding potential is turned into an operative sugar code.

In the first place, the abundant presence of freely rotatable hydroxyl groups makes carbohydrates appropriate for bidirectional acceptor/donor hydrogen (H)-bonds, which

often include coordination of metal ions or water-mediated bridging. The sp^3-hybridised oxygen atom can participate as an acceptor in two H-bonds, while the proton can act as donor. Carbohydrate-binding sites profusely contain polar residues with planar side-chains (Asn, Asp, Gln, Glu, Arg, His and Lys), also able to rotate and participate in hydrogen-bonding (Solís *et al.*, 2009). Usually, complex formation involves a few (key) hydroxyl groups of the sugar directly engaged in strong hydrogen bonds, while others participate in weaker hydrogen bonds or remain exposed to the solvent and do not interact with the protein (Solís *et al.*, 2009). The contribution of each group to the binding can be estimated through chemical mapping studies in solution by the use of engineered ligands (Solís and Díaz-Mauriño, 1997).

As significant example, ConA specificity for mannose/glucose-containing oligosaccharides involves reliable sensing through topologically fixed bidentate H-bonding. To this end, the equatorial position of the 4-hydroxyl group in mannose/glucose is critical by contributing to the establishment of appropriate H-bonds. Worth mentioning, the organisation of the binding site is assisted by a structural Ca^{2+} ion, required for ConA activity, which stabilises a *cis*-peptide bond between Ala207-Asp208 (Bouckaert *et al.*, 2000) and makes the epimer distinction possible (Gabius, 2009; Gabius *et al.*, 2011), as depicted in Fig. 1.6 (*a*).

In addition to H-bonding, other types of interactions contribute to the affinity for an epitope in solution. For instance, positively charged Lys and Arg residues are prominently involved in the binding of negatively charged sugars (containing carboxylate, sulphate or phosphate groups), evidencing that electrostatic interactions dominate recognition (Lasky, 1995; Solís *et al.*, 1998).

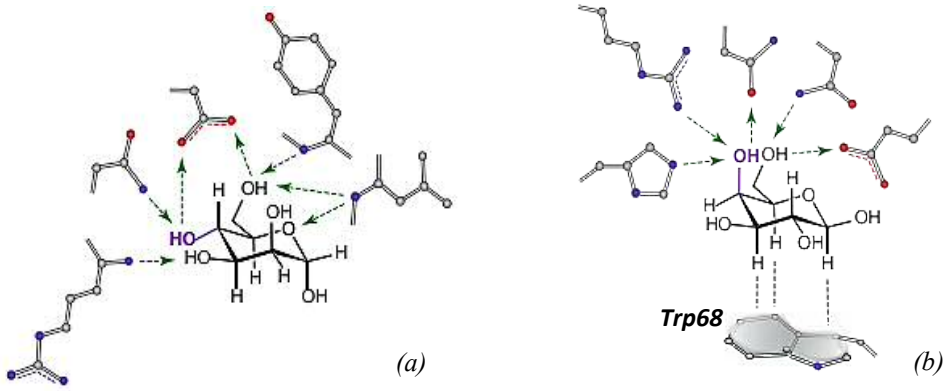


Fig. 1.6 Epimer selection by two types of lectins. *(a)* Basic atomic features of α-D-mannose recognition by ConA: Importance of the equatorial 4' OH and 6' OH in fixed bidentate H-bond formation. *(b)* Basic atomic features of β-D-galactose recognition by hGal-1: CH-π interaction established between the sugar CH-patch and the delocalised π-electron cloud of tryptophane-68. Adapted from (Gabius *et al.*, 2011).

Also, the set of C-H bonds permits hydrophobic interactions, while other non-polar groups such as the methyl moiety of acetamido groups may be involved in van der Waals contacts. Furthermore, sugar-binding sites often contain a suitably positioned aromatic residue (Trp, Tyr and, less commonly, Phe) whose delocalised π-electron cloud helps in the establishment of stacking interactions with the C-H patches of positively polarised character, as depicted in Fig. 1.6 (*b*) for human galectin-1 (hGal-1) (López-Lucendo *et al.*, 2004). The orientation of the aromatic side-chain is frequently a significant determinant of specificity, by impeding the binding of incorrect epimers (Lemieux, 1989). Finally, solvophobic effects might also be significant, since sugars perturb solvent structure in their vicinity, with diverse enthalpic/entropic consequences. In addition, complementarity of shapes and different degrees of sugar engagement in lectins' shallow grooves or deep pockets is definitely a strategic feature (Gabius *et al.*, 2011).

In any disaccharide, the relative orientation of the two pyranose rings can be easily characterised by two or three (in the case of the 1-6 bond) torsion angles of the glycosidic bond, Φ, Ψ and Ω, as depicted in Fig. 1.7. In analogy to Ramachandran's two-dimensional Φ and Ψ conformational maps for dipeptides, similar diagrams can be established for disaccharides, featuring a simplified projection of the potential energy as a function of the given coordinates Φ, Ψ or/and Ω (Kožár *et al.*, 2009). In fact, some torsion angles would give conformations that are sterically hindered, while others would maximise hydrogen bonding with a given receptor. Unlike peptides, only a few conformers are energetically privileged in glycans, defining the minima in conformational energy maps. Besides, there are demonstrated cases of selection of a minor conformer of the ligand from an equilibrium mixture in solution, such as the recognition of C-lactose by bovine galectin-1 (Asensio *et al.*, 1999). Thus, the lectin presents a kind of molecular "lock" into which a certain "key" – or group of "keys" – can fit by fulfilling certain requirements as for spatial coordination and thermodynamic selectivity. It is thus clear that besides the specific protein–carbohydrate contacts other factors govern the lectin–sugar interaction.

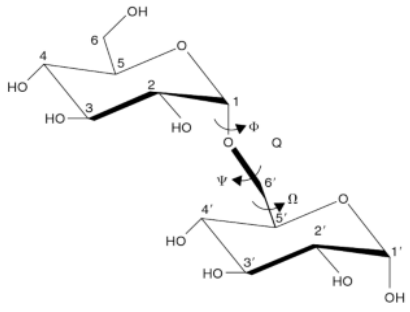


Fig. 1.7 Schematic representation of isomaltose (D-Glc-*α*(1-6)D-Glc) along with the torsion angles that define the conformations at the glycosidic linkage between the two contiguous residues: $\Phi = O_5\text{-}C_1\text{-}O_1\text{-}C_{4'}$, $\Psi = C_1\text{-}O_1\text{-}C_{4'}\text{-}C_{5'}$, $\Omega = O_1\text{-}C_{6'}\text{-}C_{5'}\text{-}O_{5'}$.

The binding affinity (or the Gibbs free energy of binding, ΔG^o) is the outcome of the enthalpy/entropy gains and penalties derived from the different events associated with the recognition process (lectin, sugar, solvent). In detail, the net enthalpy change encompasses the favourable contribution of protein–sugar contacts and of the new water–water H-bonds formed upon desolvation of lectin and sugar, as well as the unfavourable contribution of the rupture of protein–water and sugar–water interactions. The net entropy change includes the entropy gain of water molecules ordered over non-polar regions or directly bound to the protein (or sugar groups) when they return to the bulk, along with an entropy reduction of distressed water molecules that surrounded protein polar surfaces or ligands (Solís *et al.*, 2009). Besides, other events imposing entropic penalties should be considered, such as restrictions in rotation of the ligands or the loss of conformational freedom upon complex formation. In fact, even if the sugar-binding site might be pre-shaped and not undergo significant conformational changes, any alteration of the flexibility and internal mobility of the whole macromolecule could affect the binding thermodynamics (López-Lucendo *et al.*, 2004; Nesmelova *et al.*, 2010).

All things considered, the thermodynamic analysis of protein-carbohydrate interactions typically reveals enthalpy-entropy compensation, binding events being mostly enthalpically driven. Even though changes may be substantial, the net Gibbs energy (ΔG^o) is relatively small for mono- or disaccharides, with affinities in the millimolar range (Solís *et al.*, 2009). Given this relatively low affinity, extended binding sites have been established as another mechanism of affinity modulation (Rini, 1995), efficiently operating at the level of small lectin domains and increasing the affinity to the micromolar range. The 43-amino acid lectin hevein, for instance, reversibly binds to chitin, a $\beta(1\rightarrow4)$-linked *N*-acetyl-glucosamine (GlcNAc) polysaccharide. When testing oligomers of GlcNAc with increasing chain length, the ΔG^o values determined by isothermal titration calorimetry increase from -3.8 kcal/mol for the dimer up to -7.8 kcal/mol for the pentamer (Asensio *et al.*, 2000). The contact (sub)site for the activity inhibitory hapten (mono- or disaccharide) is embedded into a more complex sugar contact region – even able to accommodate oligomers. Plant defence proteins, class I chitinases (Collinge *et al.*, 1993), and chitinase-like proteins, such as the human cartilage glycoprotein-39 (Fusetti *et al.*, 2003), possess a series of up to six subsites that ensure tight binding of chitin fragments and contribute to the overall defence function. Thermodynamically, the process is driven by a favourable enthalpy through the establishment of additional contacts that are associated with a smaller

entropic penalty, as the overall motion and conformational freedom of the oligosaccharide as already been restricted, at least in part, upon binding to the primary subsite (Solís *et al.*, 2009).

Apart from one-chain oligosaccharide accommodation, glycan branching may also have an important effect on affinity, as remarkably seen in the glycoside cluster effect – a valency increase in ligands translates into a geometrical increase in affinity (André *et al.*, 2006; Lee and Lee, 2000). The signalling and regulatory potential of natural cluster formation of such lectin-reactive epitopes is now starting to be unveiled. High local epitope density may be attributable to branches of the same glycan chain, but also to the vicinity of neighbouring glycan chains (Gabius *et al.*, 2011). Overall, a high glycan density in three dimensions, as happens in the glycocalyx or the extracellular matrix, increases the propensity of lectin association (Dam and Brewer, 2010).

On the other hand, evolutionary diversification has generated divergences in the topology and/or spatial arrangement of the CRDs in different lectins, what is termed *subunit multivalency* (Rini, 1995). Clusters of binding sites can be created by either interchain disulphide bond association – as in the illustrative case of the trimeric, endogenous C-type lectin SP-D (Fig. 1.5) (Crouch *et al.*, 1994), or by tandem-repeat display in the same peptidic chain – as in the case of galectins-4, -8 and -9 (Cooper, 2002). Besides spatial vicinity, a lower entropy cost for binding to the second binding site, due to low flexibility and restricted mobility of the ligand, favours a stronger complex formation (Solís *et al.*, 2009). The spatially favoured binding steps towards saturation may exhibit particularly high affinities, down to the nanomolar range. This feature is relevant for understanding binding preferences on the cellular or pathogen surface, in terms of multivalency of both ligands and CRDs (Dam and Brewer, 2010).

1.5 Glycosylation patterns as markers for self and non-self recognition

Since the term glycobiology was introduced (Rademacher *et al.*, 1988), the role of glycan recognition by endogenous lectins in intercellular communication and within the cell itself has become increasingly considered. As described in section 1.2, a toolbox of sequentially and competitively acting factors allows both temporal and developmental regulation of the glycosylation patterns. Almost 2% of the mammalian genome, evolutionarily conserved, is dedicated to the codification for a portfolio of enzymes and transporters involved in glycosylation, of relevance to sustaining normal physiology (Ohtsubo and Marth, 2006).

Product of the coordinated action of the glycosylation machinery, the cell surface glycans constitute a hallmark for immune cell activation, differentiation and homeostasis (Marth and Grewal, 2008). Alterations and aberrant forms of the involved enzymes may cause disruptions in developmental signal-transduction pathways (Fortini, 2001) or even compromise the viability of embryonic stem cells and mammalian ontogeny (Shafi *et al.*, 2000).

The level of specialisation and regulation is high, often making glycans functionally redundant outside particular biological contexts (Lowe and Marth, 2003). For example, deletion of the *ST6Gal-I* locus (α2,6-sialyltransferase) results in a profound immunosuppression, having little effect elsewhere (Hennet *et al.*, 1998). On the other hand, ablation of the Golgi α-mannosidase II chronically activates the innate immune system inducing in mice an autoimmune disease similar to human systemic lupus erythematosus (Chui *et al.*, 2001; Green *et al.*, 2007). N-glycans produced under this enzyme deficiency bear mannose structures, ligands for innate immunity lectins, and provoke a disturbance of a fundamental defence mechanism. It should be acknowledged that mammalian N-glycan processing safeguards against the generation of an endogenous immunological signal of non-self. Thus, glycan regulation draws the narrow line between tolerance and immune activation.

Despite their ubiquitous extracellular presence and claimed pivotal role in immunity, sugars were described as T-cell-independent antigens up until recently. In particular, pure carbohydrates had previously shown relatively poor immunogenicity when used for immunisation against bacterial pathogens (Beuvery *et al.*, 1982). The development of a new generation of vaccines was consequently pursued, involving carbohydrate conjugation to a carrier protein and forcing the adaptive immune system to generate a thymus-dependent response against the carbohydrate moiety (Kasper *et al.*, 1996). This strategy effectively led to the production of protective antibodies against the bacteria from which the carbohydrates originated, as in the commercially available vaccines against *Haemophilus influenzae* type B (Kelly *et al.*, 2004) and *Streptococcus pneumoniae* (Darkes and Plosker, 2002). Simultaneously, pathogen-associated and naturally occurring carbohydrate structures, even including protein-free polysaccharides, were found to be presented by antigen-presenting cells (Tzianabos *et al.*, 2000; Zeng *et al.*, 1997) and capable of activating T-helper cells via T-cell receptor recognition (Cobb and Kasper, 2005). These advances illustrated the indisputable relevance of the recognition of non-self

glycopatterns in the induction of immunologic memory and immunoglobulin class-switching.

In contrast to the vast repertoire of antigen-specific immunoglobulins that the adaptive immune system holds, the innate immune system recognizes microorganisms via a limited number of germline-encoded pattern-recognition receptors (PRRs), which are constitutively expressed. PRRs target essential molecules associated to microbial survival or virulence, evolutionarily conserved (Akira *et al.*, 2006), unique to microbes, and invariant among microorganisms of a given class (Medzhitov and Janeway, 2002), such as endotoxin (lipid A) for Gram-negative bacteria or lipoteichoic acids for Gram-positive. A subset of PRRs, including Toll-like receptors and endogenous lectins, bridges innate with adaptive immunity by detecting explicit non-self glycosylation patterns and, subsequently, activating signalling pathways or interfering in inflammation (Barton and Medzhitov, 2003). At least three classes of lectins, C-type lectins, siglecs and galectins (see section 1.3 for classification) have been long described as PRRs in the literature (van Kooyk and Rabinovich, 2008). Some of their essential functions in innate immunity are given through examples in Table 1.2.

Table 1.2 Examples of endogenous lectins involved in innate immunity as PRRs

Lectin	Type	Specificity	Functions
Mannose-binding lectin	C	Terminal Man /GlcNAc	Activation of the complement lectin-pathway
Surfactant proteins (-A, -D)	C	Man, maltose	Increase bacterial membrane permeability, opsonise pathogens, regulate production of inflammatory mediators
DC-SIGN	C	Internal/ terminal Man	Pathogen recognition, antigen presentation or internalisation, interaction with T-cells
Dectin-1	C	β1,3 and 1,6 glucans	Anti-fungal defence by induction of TNF-α, regulation of T-cell proliferation
Ficolins (L-, H-)	Fibrinogen-like	GlcNAc/ GalNAc	Activation of the complement lectin-pathway
Galectins (1-15)	S	β-Gal	Increase bacterial membrane permeability and modulate immune activation and inflammation
Serum amyloid P component	Pentraxin	Gal (4,6-pyruvate acetal)	Opsonises pathogens, promotes their complement-mediated clearance
Siglecs	I	Sialic acids	Increase pathogen uptake by macrophages and modulate immunogenic or tolerogenic programs of antigen-presenting cells

References: (Bottazzi *et al.*, 2016; Crocker *et al.*, 2007; Crouch *et al.*, 2000; Fujita, 2002; McGreal *et al.*, 2004; Rabinovich *et al.*, 2012; Stowell *et al.*, 2014; Wright).

The accessibility and density of glycans on the pathogen surface strongly contributes to their recognition by lectins (Reid *et al.*, 2010). Interestingly, molecular mimicry is exploited by pathogens that disguise by masking their patterns with host-self antigens (Moran *et al.*, 1996). Most nasopharyngeal pathogens, including nontypeable *H. influenzae* present on their surfaces sialic acid or digalactose units that function as anti-recognition molecules, allowing the microbe to masquerade as "self" (Vimr and Lichtensteiger, 2002) and to avoid complement-mediated bacteriolysis (Fujita, 2002; Hood *et al.*, 1999; Weiser and Pan, 1998).

On the other hand, in addition to acting as self-recognition signals, mammalian glycans, densely coating every cell, also serve as docking sites for pathogen receptors, as described above (Karlsson, 1995; Sharon, 1996). Pathogens could thus act as an exogenous selection pressure upon their long-lived hosts obliging them to shift their glycopatterns in order to evade infection, without compromising their own survival (Gagneux and Varki, 1999). In this setting, glycosylation patterns seem to be trapped in never-ending cycles of evolutionary "Red Queen" effects. The name of this evolutionary concept derives from the Red Queen's comment to Alice in *Through the Looking Glass* that "it takes all the running you can do, to stay in the same place" (Van Valen, 1974). Hence, the remarkably well conserved mammalian glycans are prone to undergo rapid, even constant, evolutionary change to guarantee the organism survival (Gagneux and Varki, 1999), causing inter-individual deviations and also leading to the emergence of distinct susceptibility traits towards infectious agents (Varki, 2006).

The glycosylation patterns can also be altered in disease, cancer being the most prominent example. Tumour cells may display an aberrant glycosylation profile, at both structural and expression levels, compared with their non-transformed counterparts (Pinho and Reis, 2015). Two principal mechanisms underlying the tumour-associated alterations of carbohydrate structures were postulated in the so-called *incomplete* synthesis of truncated structures (due to pathway impairment) and *neo-synthesis* (upon gene induction) processes (Hakomori, 2002). In the ´80s, malignant transformation was also found to be often accompanied by the expression of oncofetal antigens decorating proteins and sphingolipids (Feizi, 1985; Hakomori, 1985), which allow neoplastic cells to usurp many of the events naturally occurring in development, i.e. cell adhesion and motility. At that time, further insights into the unique repertoire of glycans expressed on tumour cells emerged from their recognition by a variety of plant lectins of known binding specificity (Raedler *et al.*, 1988).

Evidently, elucidation of the glycosylation patterns in normal and tumour cells, as well as engineered living cells (Hudak *et al.*, 2014) will help to correlate glycan structure and function. In a recent study, remodelling the sialylation status of the glycocalyx of distinct cancer cell types affected their susceptibility to innate natural killer cell cytotoxicity through the recruitment of the inhibitory sialic acid-binding immunoglobulin-like lectin 7 (Siglec-7) (Hudak *et al.*, 2014). Although the tumour-associated carbohydrate antigens represent altered self-epitopes rendering cancer cells mildly antigenic, they are rarely immunogenic or capable of eliciting strong humoral responses (Amon *et al.*, 2014; Dube and Bertozzi, 2005) unless presented on an alternative scaffold, such as virus-like particles (Yin and Huang, 2015). A comprehensive study of interactions of endogenous lectins involved in the innate immunity, including selectins and other C-type lectins (Kohler *et al.*, 2009), galectins (Rabinovich and Toscano, 2009) or Siglecs (Macauley *et al.*, 2014), will also shed light to the mechanisms that contribute to health and disease (Macauley and Paulson, 2014; Solís *et al.*, 2015). Non-invasive screening of cancer-associated glycopatterns remains an aim hard to achieve: glycoproteins showing aberrant profiles are clinically utilised as serological biomarkers in cancer patients, but their use in early diagnosis is precluded because of their concurrent expression in benign diseases (Pinho and Reis, 2015).

Only recently, extracellular vesicles of both prokaryotic and eukaryotic origin were found to be capable of transferring molecular patterns, metabolites or effectors (siRNAs) in a horizontal manner, from one cell to another (Senfter and Mader, 2015). Extracellular membrane vesicles are heterogeneous structures and can be classified into several groups based on their size, antigenic features and mechanism of cellular release (Thery *et al.*, 2009), as described in Table 1.3 and Fig. 1.8. The two best characterised categories include exosomes (EXO) and microvesicles/microparticles (MVs), whose release by various cell

Table 1.3 Key features of extracellular vesicle populations

	Exosomes	**Microvesicles**	**Apoptotic vesicles**
Size range	50-150 nm	~100-500 nm	>800 nm
Mechanism of generation	Multivesicular bodies exocytosis	Plasmatic membrane budding and/or blebbing	Plasmatic membrane blebbing during apoptosis
Detection	Electron microscopy, western blotting, bead-coupled flow cytometry	Flow cytometry and capture-based assays	Flow cytometry
Markers	Annexin V, CD63, CD81, CD9, etc.	Annexin V, tissue factor and cell-specific markers	Annexin V, DNA content

Adapted from (György *et al.*, 2011b; Thery *et al.*, 2009).

types is enhanced during cell activation and apoptosis (György *et al.*, 2011b; Raposo and Stoorvogel, 2013).

Exosomes, membrane enclosed vesicles of 50-150 nm diameter, are singularly assembled and secreted following an orchestrated procedure (Trajkovic *et al.*, 2008), and are thought to share glycosylation patterns with their parental cell. They carry aberrant glycoforms when excreted by malignant cells and this property could turn them into functional, pluripotent clinical biomarkers for non-invasive diagnosis. Indeed, in a recent study, the proteoglycan glypican 1 was used to distinguish circulating cancer exosomes from benign vesicle carriers, thus allowing early pancreatic cancer diagnosis (Melo *et al.*, 2015). These findings open new horizons for the search of glyco-biomarkers in the exosomes field, even though straight-forward and time-effective approaches for glycan profiling are currently missing. In parallel, comprehensive glycosylation pattern analyses of distinct vesicle populations are expected to answer questions as for their origin and natural role, and pave the way for cell-targeting applications.

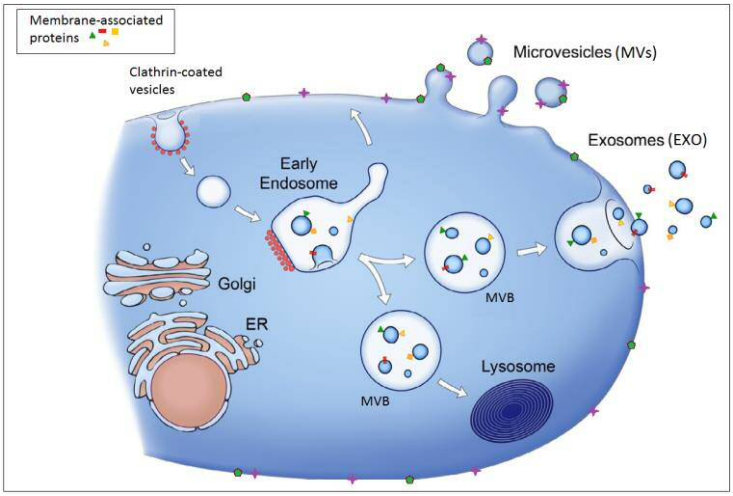
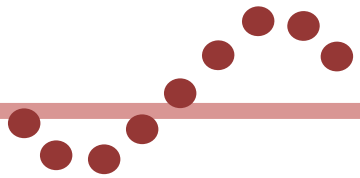


Fig. 1.8 Cellular release of microvesicles and exosomes. MVs bud directly from the plasma membrane. EXO are formed by blebbing into early endosomes or multivesicular bodies (MVBs), then released by fusion of the latter with the cellular membrane. In some cases, they may also fuse with lysosomes. The point of divergence is suspected to lie on distinct early endosomes feeding or labelling. Adapted from (Raposo and Stoorvogel, 2013).



OBJECTIVES

The establishment of structure-function correlations for glycans, known as functional glycomics, has reshaped post-genomics biology and is undoubtedly a complex task. Even though comprehensive glycan sequencing may not be always necessary for the attribution of functions, a multitude of other features still have to be taken into account: clustering of saccharide patches decorating the cellular surface, epitope microheterogeneity, or multivalency to overcome low intrinsic affinities. To achieve advanced glycan profiling and functional characterisation, the merge of the use of lectins as unique pattern-reading tools with a combination of well-established analytical methodologies seems compulsory.

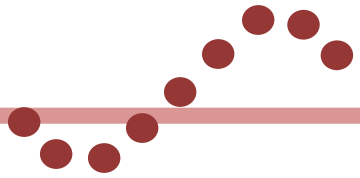
The use of lectins for decades in the glycophenotyping of glycoproteins, pathogens and tissues (Gabius *et al.*, 2011) is only recently extending to undifferentiated cells (Hamouda *et al.*, 2013) and extracellular vesicles (Gerlach *et al.*, 2013; Gomes *et al.*, 2015). In the case of respiratory pathogens, the virulence degree of serological types distinguished by their surface carbohydrate structures appears to be associated with their recognition, or lack of recognition, by lectins of the innate immune system (Sahly *et al.*, 2008). Thus, pathogen glycoprofiling and identification of epitopes recognised by endogenous lectins could pave the way for establishing virulence correlations.

To facilitate the task of identification of lectin ligands by high-throughput screening, glycan microarray platforms have been extensively used. Yet, a main limitation is that they are largely built with mammalian-like glycan libraries, whereas bacteria are known to display distinct monosaccharide units (Adibekian *et al.*, 2011). More focused libraries that incorporate synthetic or natural bacterial glycans have also been developed (Stowell *et al.*, 2014). Still, glycans' presentation and density in the array may substantially differ from their natural arrangement, a factor reported to affect recognition, particularly in the case of weak binders (Godula and Bertozzi, 2012; Wang *et al.*, 2014). Consequently, the real accessibility and density of each glycan on the target surface, as e.g. bacteria, extracellular vesicles, or entire cells cannot be assessed. Furthermore, a possible binding event arising from a simultaneous recognition of different, neighbouring surface molecules would pass unnoticed in a carbohydrate microarray setup, where binding to individual glycans is tested. To overcome these limitations, a novel application of the microarray technology has been recently developed based on the generation of bacteria microarrays, using *Klebsiella*

pneumoniae, a Gram-negative, capsulated and LPS-containing bacteria, as model pathogen (Campanero-Rhodes *et al.*, 2015).

The core objective of this PhD Thesis is the design and development of novel platforms for the study of lectin-carbohydrate interactions directly on the probe surface, thus preserving the carbohydrate epitopes´ natural assembly and vicinity. Specific objectives are:

- Development of designer' microarray and QCM strategies for the analysis of pathogen surface glycosylation and exploration of the recognition by lectins of the innate immune system. Nontypeable *Haemophilus influenzae* (NTHi) was selected as relevant example of nonencapsulated opportunistic pathogen bearing a lipopolysaccharide without O-antigen and, therefore, a limited number of carbohydrate structures at its surface.
- Design and establishment of extracellular vesicle microarrays as sensitive tools for profiling the glycosylation patterns of exosomes and microvesicles of different origin, in search of malignancy biomarkers.
- Setting-up of microarrays for evaluation of the functionality of bacterial adhesins on live bacteria and the screening of carbohydrate-bearing nanosystems of potential use in anti-adhesive strategies.



Materials & Methods

3.1 Bacteria culture, fixation and staining

Nontypeable *Haemophilus influenzae* (NTHi) strains (Table 3.1) were seeded in a chocolate-agar plate (Biomérieux) directly from the stock conserved at -80 ºC, and grown overnight at 37 ºC in the presence of 5% CO_2. Each strain was inoculated (1 to 5 colonies) in 40 ml of brain-heart infusion medium supplemented with 10 µg/ml hemin and 10 µg/ml β-nicotinamide adenine dinucleotide (β-NAD), hereafter referred to as sBHI, and incubated under shaking at 180 rpm, 37 ºC and 5% CO_2 for 12 h. Cultures having reached the stationary phase were then diluted (1:10) in 40 ml fresh sBHI and incubated, at 37 ºC and 5% CO_2, till they reached the exponential growth phase (OD_{600} = 0.3-0.4). The viability of each culture was tested by serial dilution and plating on sBHI-agar. Subsequently, 15 ml of each culture was collected and centrifuged at 4,000×g for 15 min at 4 ºC. Pellets were resuspended and adjusted to OD_{600} = 1 in 5 mM sodium phosphate buffer, pH 7.2, containing 0.2 M NaCl (PBS).

For fixation, a previously established protocol was followed (Amann *et al.*, 1990) with slight modifications. In detail, 4 ml of each adjusted culture was centrifuged at 4,000×g for 15 min at 4 ºC, subsequently resuspended in 750 µl PBS plus 250 µl cold 4% paraformaldehyde aqueous solution and incubated for 24 h at 4 ºC. Each preparation was centrifuged at 4,000×g for 10 min and washed twice with 500 µl Tris-HCl buffer saline (10 mM Tris, pH 7.8, containing 150 mM NaCl, TBS), resuspended in TBS, then conserved at 4 ºC until used. Culture and fixation of NTHi was carried out at the laboratory of Dr. Junkal Garmendia (Instituto de Agrobiotecnología, CSIC-UPNa-Gobierno de Navarra, Mutilva, Spain).

The final preparations of fixed bacteria were labelled with SYTO® 13 green fluorescent nucleic acid stain (Life Technologies). SYTO® dyes belong to the class of substituted asymmetrical cyanine dyes, thought to have multiple nucleic acid binding modes, including intercalation and charge interactions with the phosphate backbone in the minor groove (Eischeid, 2011; Tárnok, 2008). Detailed structural information is generally not available for commercial dyes. SYTO® dyes have low intrinsic fluorescence with quantum yields (QY) typically <0.01 when not bound to nucleic acids or stand free in solution. When bound to DNA/RNA, SYTO® 13 has a QY of 0.4. Since its excitation maximum is 488 nm and its emission maximum is 509 nm, it can be excited by using a blue (488 nm) laser.

Table 3.1 Nontypeable *H. influenzae* strains and mutants[a] used in this Thesis

Strain	Description	Phenotypic traits	Clinical data	Source	Reference or source
NTHi375	Clinical isolate	Whole genome sequence, methylome, and LOS structure available	Otitis media	Middle ear	(Mell *et al.*, 2014)
86-028NP	Clinical isolate	Whole genome sequence available	Otitis media	Nasopharynx	(Harrison *et al.*, 2005)
NTHi398	Clinical isolate	Whole genome sequence available	COPD, exacerbation	Sputum	(De Chiara *et al.*, 2014)
NTHi1566	Clinical isolate	Whole genome sequence available	COPD, exacerbation	Sputum	(De Chiara *et al.*, 2014)
NTHiP131	Clinical isolate	Genome not sequenced. Distinct PFGE profile	Pediatric healthy carrier	Nasopharyngeal swab	(Kalograiaki *et al.*, 2016)
NTHiP133	Clinical isolate	Genome not sequenced. Distinct PFGE profile	Pediatric healthy carrier	Nasopharyngeal swab	(Kalograiaki *et al.*, 2016)
Rd KW20	Laboratory strain	Whole genome sequence available. Capsule-deficient, serotype d			(Fleischmann *et al.*, 1995)
Rd *hmw1*$_{strain12}$	Laboratory derivative	Rd strain expressing *hmw1* locus of NTHi12 (R2846) – *hmw1A*, *hmw1B*, *hmw1C*			(Grass *et al.*, 2003)
NTHi375Δ*opsX*	*opsX::ermC*, Erm^r	LOS molecule is truncated at Kdo level, it does not contain tri-Hep backbone			(Morey *et al.*, 2013)
NTHi375Δ*lgtF*	*lgtF::ermC*, Erm^r	LOS molecule lacks oligosaccharide extensions at HepI			(Morey *et al.*, 2013)
NTHi375Δ*lpsA*	*lpsA::ermC*, Erm^r	LOS molecule lacks oligosaccharide extensions at HepIII			(Morey *et al.*, 2013)
NTHi375Δ*lgtF*Δ*lpsA*	*lgtF::ermC*, *lpsA::km*, Erm^r Km^r	LOS molecule lacks oligosaccharide extensions at HepI and at HepIII			(Morey *et al.*, 2013)
NTHi375Δ*lic2A*	*lic2A::ermC*, Erm^r	LOS molecule lacks digalactose			(Morey *et al.*, 2013)
NTHi375Δ*siaB*	*siaB::ermC*, Erm^r	LOS molecule lacks sialic acid			(Morey *et al.*, 2013)
NTHi375Δ*ompP5*	*OmpP5::ermC*, Erm^r	Porin P5 is not expressed			(Euba *et al.*, 2015)

[a] For mutant designation, the symbol Δ is used to indicate disruption of the respective gene

For optimal staining, 1 µl of SYTO® 13 dye (5 mM stock solution in DMSO) was added to the bacterial suspension of a total volume of 500 µl, giving a final concentration of 10 µM. Bacterial suspensions were vortexed upon addition of the dye and incubated for 10 min at 20 ºC. Then, they were centrifuged at 4,000×g for 15 min and the pellets were washed twice with fresh TBS and conserved at -20 ºC until used. When manipulated mostly in the dark, no loss of fluorescent signal was observed, as evaluated after 2, 4, 6 and 12 months of storage.

Escherichia coli strain UTI89 was kindly provided by Prof. Jean-Jacques Letesson (Laboratory of Infection, Cell & Host, Research Unit in Microorganism Biology, Université de Namur – FUNDP, Namur, Belgium). The above-described culture protocol was followed with the difference that the culture medium used was enriched lysogeny broth (LB2YT) containing 8 g of tryptone, 5 g of yeast extract and 5 g of NaCl per 500 ml. The concentration of *E.coli* was estimated by correlating turbidity measurements with number of colonies obtained after plating serial dilutions (0.88-1.13×10^9 colony-forming units (cfu) per ml. Fifteen ml of culture in the desired growth phase (either exponential or stationary) was collected and centrifuged at 4,000×g for 15 min at 4 ºC. Pellets were washed and resuspended in 5 ml of 10 mM HEPES buffer, pH 7.4, containing 150 mM NaCl and 10 mM $CaCl_2$ (HBSC). For staining, 10 µl of SYTO® 62 red fluorescent nucleic acid stain (5 mM solution in DMSO, Life Technologies) was added to the bacterial suspension, giving a final concentration of 10 µM. SYTO® 62 when bound to DNA/RNA has a QY of 0.27, excitation maximum at 652 nm and emission maximum at 676 nm, and can be excited by using a red (635 nm) laser. Following 10 min incubation in the dark, the bacteria were centrifuged and washed, then resuspended in HBSC buffer containing 0.1% Tween®20 and used directly.

Labelling efficiency with SYTO® dyes was evaluated by fluorescence spectroscopy using a FluoroMax®-4 (Horiba Jobin Yvon) spectrometer and Suprasil®-quartz (QS) micro-cuvettes with PTFE stopper (lightpath: 5 mm × 5 mm, max. volume: 850 µl, Hellma Analytics). Emission spectra were acquired for bacterial suspensions of turbidity OD_{600} ~0.5 with the assistance of FluorEssence™ software (Horiba Jobin Yvon), using as excitation wavelength 488 nm and 652 nm for SYTO® 13 and 62, respectively. Increment was set at 1 nm and the excitation and emission slit at 2 nm. Optimal integration time resulting in an acceptable signal-to-noise (S/N) ratio was determined to be 0.5 sec. Both Signal (counts per second, cps) and Reference

(microAmps) detectors were activated in current-acquisition mode and the S/R ratio was registered. The ratio acquired resulted in signal counts independent of changes in the light-source intensity through time.

For SYTO® 13-stained bacteria, the maximum of the emission spectra was at 507 nm and for SYTO® 62 was at 676 nm. Each spectrum was acquired twice and the mean value at 507/676 nm for each strain was correlated with the turbidity of the sample and colony counts before bacteria fixation, to compare labelling efficiency among strains.

3.2 Isolation of nontypeable *Haemophilus influenzae* lipooligosaccharide

Lipooligosaccharide (LOS) molecules of selected NTHi strains were extracted using a phenol-water extraction protocol (Johnson and Perry, 1976), followed by enzymatic treatment to eliminate contaminating protein and nucleic acids. In detail, NTHi strains grown on chocolate-agar plates were inoculated (3 to 5 colonies) in 400 ml sBHI and incubated for 12 h under shaking at 180 rpm, 37 °C and 5% CO_2. Viability was determined by serial dilution and plating on sBHI-agar, yielding a comparable number of bacterial cfu per ml for all cultures. Bacterial pellets were collected by centrifugation at 6,000×g for 15 min at 4 ºC, resuspended in water (33.33 ml water/g of dry pellet), and thoroughly mixed with an equal volume of equilibrated phenol solution (Sigma-Aldrich) at 66 ºC. Following centrifugation at 8,000×g for 15 min at 4 ºC, the phenol layer was discarded and the LOS, present in the aqueous layer, was precipitated by adding four volumes of methanol containing 1% of methanol saturated with sodium acetate, and freezing at -20 ºC for at least 12 h. The precipitate was recovered by centrifugation, resuspended in 5 ml of water and exhaustively dialyzed against high-purity water (milli-Q®). The dialysate was then centrifuged at 100,000×g for 6 h at 4 ºC and the resulting pellet was resuspended in 2 ml of water and lyophilised. For elimination of contaminating nucleic acids, the lyophilisate was resuspended at 10 mg/ml in 100 mM Tris-HCl buffer, pH 7.0, containing 0.05% (w/v) NaN_3, and digested with 50 µg/ml DNase II and RNase I (Sigma-Aldrich) for 30 min at 37 ºC. Next, protein traces were also digested thrice by incubation with 50 µg/ml of proteinase K (Sigma-Aldrich) at 55 ºC for 3 h. The purified LOS was precipitated with methanol/sodium acetate using the protocol described above, dialysed and high-speed centrifuged, and the resulting pellet was finally resuspended in 200 µl of milli-Q® water. LOS isolation and purification was carried out at the laboratory of Dr. Junkal Garmendia.

3.3 Isolation and physical characterisation of extracellular vesicles

3.3.1 Isolation of extracellular vesicles

Extracellular vesicles were isolated from human blood and from the THP-1 human acute monocytic lymphoma cell line (ATCC® TIB-202™), one of the most widely used cell lines to investigate the function and regulation of monocytes in different physiological systems. Isolation took place at the laboratory of Prof. Edit Buzás (Semmelweis University, Budapest, Hungary).

THP-1 cells were cultivated in RPMI-1640 medium (GIBCO®, Life Technologies) containing 1% L-glutamine, 10% (v/v) centrifuged foetal calf serum and 1% (w/v) penicillin/streptomycin (all acquired from Sigma-Aldrich), in 150 ml flasks at 37 ºC until they reached 3-5×10^5 cells/ml, as described previously (Osteikoetxea *et al.*, 2015). Then, they were changed to 60 ml fresh serum-free medium and left to vesiculate for 24 h ("vesicle farming"). Cell viability was confirmed by fluorescence-activated cell sorting (FACS) upon labelling with an Annexin V-FITC and propidium iodide Apoptosis Detection Kit (Sigma-Aldrich Life Science). The culture was centrifuged initially at 300×g for 10 min to pellet and discard cells (Fig. 3.1, *a*). The supernatant (S_B) was filtered by gravity through 0.8 µm filters to exclude apoptotic bodies (Fig. 3.1,

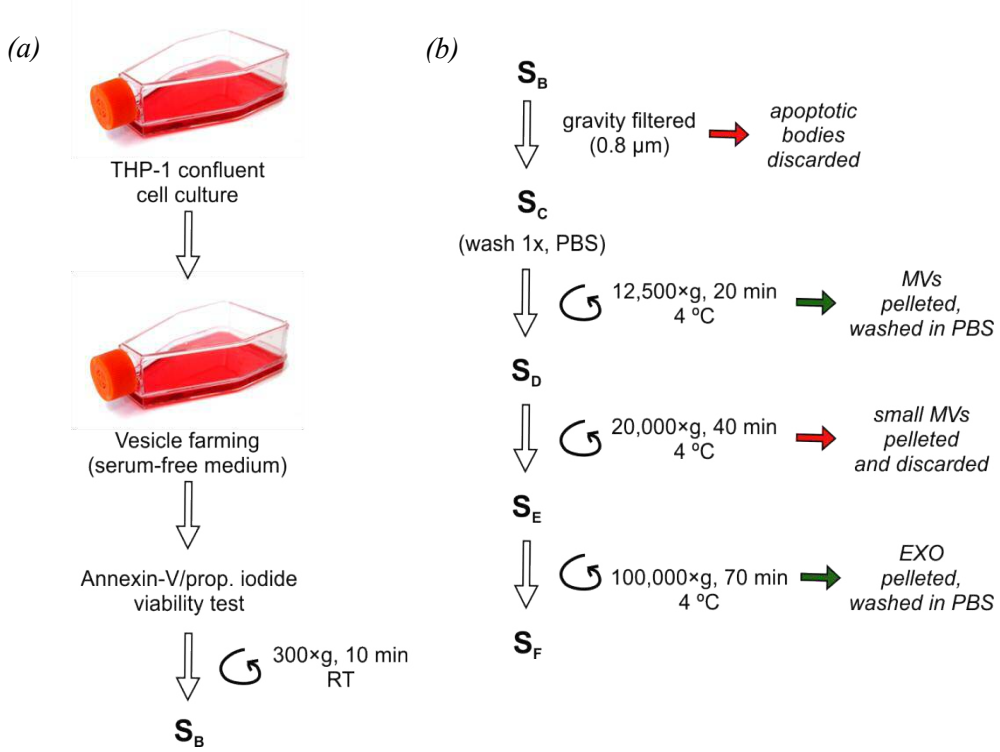


Fig. 3.1 ***(a)*** THP-1 cell culture and culture supernatant (S_B) isolation. *(b)* Fractionation of MV and EXO populations

b). The filtered supernatant (S_C) was centrifuged at 12,500×g for 20 min at 4 ºC to pellet MVs. The pellet was washed once with PBS, and MVs were resuspended in 300 µl fresh PBS by gentle pipetting. The supernatant (S_D), containing smaller MVs populations close to EXO size, was centrifuged for 40 min at 20,000×g and 4 ºC, and the pellet was discarded. The supernatant (S_E), containing EXO only, was then centrifuged at 100,000 ×g for 70 min at 4 ºC (Fig. 3.1, *b*). The supernatant (S_F) was separated and the EXO pellet was washed once with PBS, centrifuged under the same conditions and resuspended in 300 µl of PBS by gentle pipetting. All supernatants (S_B-S_F) were kept as controls.

Blood-derived extracellular vesicles were isolated as described before (György *et al.*, 2011a). In detail, 500 ml of normal human blood was collected from healthy donors and centrifuged at 250×g for 10 min to eliminate blood cells. The samples were aliquoted and stored at -20 ºC up until used. Platelets, cell debris, and apoptotic bodies were pelleted at 3,000×g for 10 min, followed by filtration of the supernatant through an 0.8 µm filter (Millipore®) by gravity. MVs and EXO were sedimented by centrifugation and resuspended in PBS as described above.

MVs and EXO samples in PBS were stored at -20 ºC until used. MVs stability was estimated to be 3 months by FACS and Annexin V-FITC staining. EXO are expected to be stable for longer due to their compact size and microdomain-lipid origin.

3.3.2 Characterisation of THP-1 derived EV preparations by Resistive Pulse Sensing

EV preparations were submitted to TRPS analysis using a qNano instrument (IZON Science, New Zealand). Serial dilutions were prepared from each EV preparation, EXO and MVs. Particle numbers were then counted for 5 min using 5 mbar pressure and NP200, NP400 and NP2000 nanopore membranes stretched between 45 and 47 mm. Voltage was set in between 0.1 and 0.25 V in order to achieve a stable 115 nA current. Particle size histograms were recorded when root mean square noise was below 12 pA, particle rate in time was linear, and at least 500 events were counted. Calibration was performed using known concentration of beads CPC100B (mode diameter: 110 nm), CPC200B (mode diameter: 203 nm) or CPC400E (mode diameter: 340 nm) and CPC2000C (mode diameter: 1,900 nm) (all from IZON) diluted 1:1,000 in 0.2 µm filtered PBS.

3.3.3 Morphological characterisation by Transmission Electron Microscopy

This experimental part was carried out by the PhD candidate at the laboratory of Dr. Ágnes Kittel (Department of Pharmacology, Institute of Experimental Medicine, Hungarian Academy of Sciences, Budapest, Hungary). Shortly after ultracenrifugation resulting in a pellet that contained extracellular vesicles, the supernatants were carefully withdrawn and the pellets fixed with 4% paraformaldehyde in PBS for 60 min at room temperature. Following washing with PBS, the preparations were post-fixed in 1% OsO4 (Taab, werAldermaston, Berks, UK) for 30 min. After rinsing the intact fixed pellets within the centrifugation tubes with distilled water, the pellets were dehydrated in graded ethanol (50, 70, 90, 96 and absolute ethanol), including block-staining with 1% uranyl-acetate in 50% ethanol for 30 min, then embedded in Taab 812 (Taab). After overnight polymerization at 60 °C and sectioning for EM, the 50-70 nm ultrathin sections were analyzed using a Hitachi 7100 electron microscope (Hitachi Ltd., Japan) equipped with a side-mounted Veleta CCD camera (lower resolution, Olympus Soft Imaging Solutions) digital camera.

3.3.4 Examination of the membrane lipid order

In order to investigate the membrane lipid bilayer order of EV preparations, EVs were stained with the membrane probe Di-4-ANEPPDHQ (Life Technologies) at 5 μM final concentration for 30 min at RT as described previously (Frescatada-Rosa *et al.*, 2014; Osteikoetxea *et al.*, 2015; Owen *et al.*, 2012). Staining was evaluated by fluorescence spectroscopy using FluoroMax®-4 spectrometer as described in section 3.1. Emission spectra (range 500-750 nm) were acquired at steady phase without previous dilution, using as excitation wavelength 488 nm. Increment was set at 1 nm, the excitation slit at 2 nm and the emission slit at 5 nm and the S/R ratio was registered. Optimal integration time resulting in an acceptable signal-to-noise (S/N) ratio was determined to be 0.5 sec. Intensities corresponding to the membrane ordered (I_{570}) or disordered (I_{630}) regions were denoted (in cps) at 570 nm and 630 nm respectively, their ratio revealing the general membrane polarization (GP), characteristic for each vesicular population. GP was calculated using the formula,

$$GP = \frac{I_{570} - G \times I_{630}}{I_{570} + G \times I_{630}}$$

where G corresponds to a compensation factor. In order to obtain the G factor, the same fluorimeter set-up was used to image the fluorescence of a similar stock dilution in DMSO and GP_{mes} values were calculated (using the previous equation with G = 1). The G factor was calculated according to the equation:

$$G = \frac{GP_{ref} + GP_{ref}\,GP_{mes} - GP_{mes} - 1}{GP_{mes} + GP_{ref}\,GP_{mes} - GP_{ref} - 1}$$

where GP_{ref} is a reference value for di-4-ANEPPDHQ in DMSO, here fixed at -0.85 as recommended (Owen *et al.*, 2012). In our set-up, the G factor was defined as G = 0.19.

3.4 Synthesis of glyconanoparticles

Glyconanoparticles (GNPs) – nanoparticles carrying surface-tethered carbohydrate ligands – have demonstrated a great potential in biomedical imaging, therapeutics, and diagnostics. Two general strategies for functionalisation can be discerned, based on either non-covalent or covalent protocols. Covalent binding is generally preferred due to the ordered ligand attachment to the constructs and their considerably higher stability (Wang *et al.*, 2010). The coupling chemistry used here is based on the established procedure for covalent attachment of carbohydrates to solid substrates (e.g. silica nanoparticles) by using functionalised perfluorophenyl-azides (PFPAs) (Wang *et al.*, 2011a). Photo-initiated coupling chemistry does not require derivatisation of the carbohydrates. The azide (N_3^-) moiety on PFPAs can be activated by UV light and converted into a highly active nitrene, which can undergo insertion reaction into diverse CH bonds or addition reaction to C=C bonds. As the establishment of only one covalent bond is necessary for insertion, the probability of bond formation increases with the number of CH bonds present in the sugar (i.e. 57% for α-D-mannose, 74% for di-mannosides, 81% for tri-mannosides, etc).

GNPs were synthesized at Prof. Olof Ramström´s laboratory (Division of Organic Chemistry, School of Chemical Science and Engineering, Kungliga Tekniska Högskolan, Stockholm, Sweden). Silica nanoparticle cores (SNPs) were synthesised following an established protocol (Stöber *et al.*, 1968) with slight modifications. Tetraethylorthosilicate (2.8 ml) was added to 200-proof absolute ethanol (34 ml) followed by NH_4OH (35%, 2.8 ml). The reaction was allowed to proceed overnight at 20 ºC with vigorous stirring to yield a white colloidal solution. To this solution, PFPA-

silane (80 mg) prepared by adding $NH_2(CH_2)_3Si(OMe)_3$ to PFPA-N-hydroxysuccinimide, NHS, (Yan and Ren, 2004) was added, and the mixture was left to react overnight (Fig. 3.2). The resulting PFPA-SNPs were purified as previously described (Wang *et al.*, 2011b).

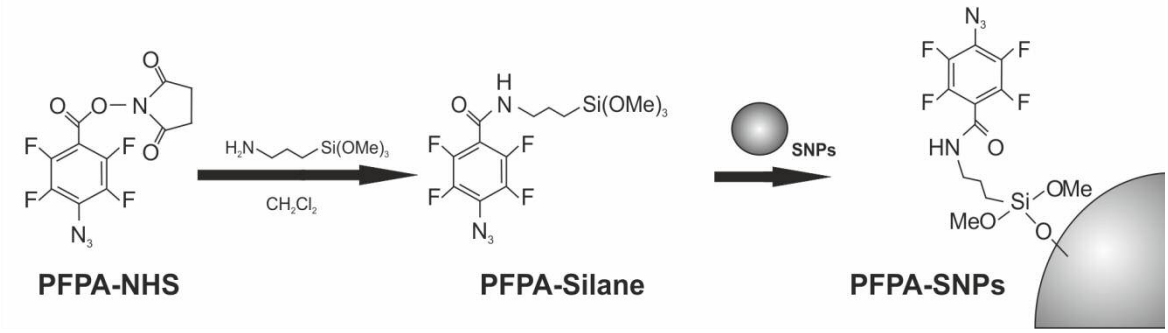


Fig. 3.2 Preparation of PFPA-SNPs.

Then, 70 mg of PFPA-SNPs was dispersed in 20 ml of acetone and sonicated for 5 min to achieve a NP-dispersion of 3.5 mg/ml. For one-pot covalent attachment (Fig. 3.3), one ml of the PFPA-SNP dispersion in acetone was placed in a flat-bottom dish and 0.5 ml of 3.5 mg/ml sugar solution in water was added. The mixture was covered with a 280 nm long-path optical filter (WG-280, Schott Glass) and irradiated with a 450 W medium-pressure mercury vapor lamp (Hanovia) for 20 min under rigorous magnetic stirring. The dispersion was then centrifuged at 4,600×g for 20 min to sediment the GNPs. These were washed once with milli-Q® water, dialysed against circulating water overnight using a 12-14 kDa-limit cellulose membrane (Spectra/Por® 4, 45 mm flat-width), centrifuged and re-suspended in 1 ml of acetone.

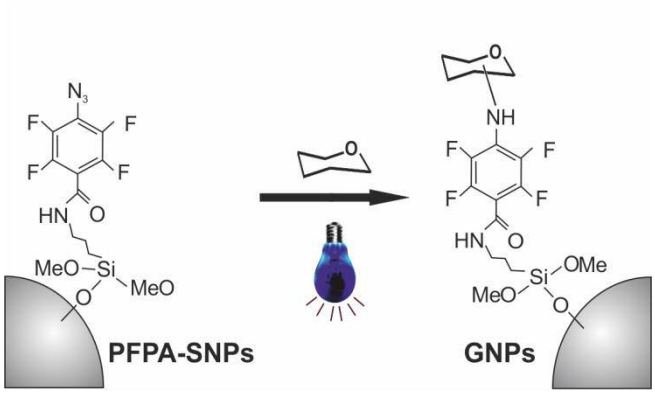


Fig. 3.3 SNP UV-assisted derivatisation. PFPA-SNPs in acetone and underivatised sugar in water were mixed under UV radiation to yield GNPs.

Upon rotary evaporation of the solvent for 15 min at 20 ºC, each sample was dried under vacuum overnight. The product was weighed and checked by infra-red (IR) spectroscopy for the presence of the azide peak (2075 cm^{-1}) (Annex I). GNPs

synthesised included galactose- (Gal, Sigma), α-methyl-mannoside- (α-Me-Man, Sigma) and methyl-3-O-(α-D-mannopyranosyl)-α-D-mannopyranoside (manα(1,3)man - α-OMe, Carbosynth) derived GNPs.

Quantitation of the immobilised ligand was carried out by hexose determination using the phenol–sulphuric acid micro-method (section 3.10.2). As opposed to the anthrone–sulphuric acid method previously employed with the same purpose (Wang *et al.*, 2011a), no background noise was produced. GNP-size was determined to be 110 nm by Trasmission Electron Microscopy, using a JEOL 100CX TE microscope operating at an accelerating bias voltage of 100 kV (Wang *et al.*, 2011c).

3.5 Designer´s microarrays

"Designer´s microarrays", a term introduced in the literature in 2006 (Palma *et al.*, 2006), permit the "designer" to incorporate any kind of probes in the array, by employing diversely functionalised slide surfaces, and to achieve multiplexed results by varying the assay conditions. In this Thesis, different microarray platforms have been developed and optimised following this concept, including protein, neoglycopeptide, polysaccharide, lipopoly-/lipooligo-saccharide, bacteria and extracellular vesicle microarrays. All platforms have been evaluated for printing quality, probe-retention capacity, stability and applicability in functional analyses.

3.5.1 Protein, neoglycopeptide and lipopoly-/lipooligo-saccharide probes

Naturally occurring glycoproteins usually contain complex and heterogeneous mixtures of oligosaccharide chains, making them outstanding detectors of lectin-carbohydrate interactions. In this Thesis, controls were carefully selected among a variety of glycoproteins with characterised glycosignatures (Fig. 3.4), to confirm carbohydrate-binding activity of the lectins tested.

Fetuin (FET, Type IV) and asialofetuin (AFET, Type I) from foetal calf serum, human fibrinogen (Fg, Type I) and ribonuclease B (RNase B, TypeXII-B) from bovine pancreas were purchased from Sigma-Aldrich. As negative control, bovine pancreatic ribonuclease A (RNase A, Type XII-A, Sigma-Aldrich) was also included as probe. RNase A is the non-glycosylated major form of the enzyme (coexisting with RNase B) and was used to detect possible non-carbohydrate-mediated interactions of the lectins

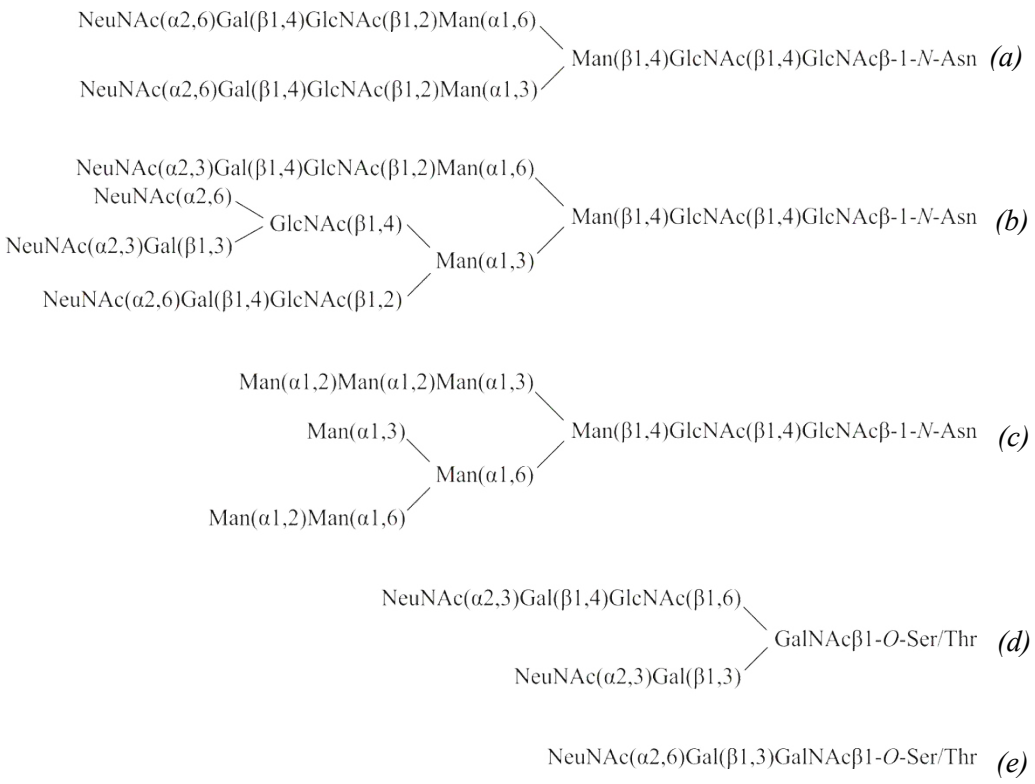


Fig. 3.4 Examples of carbohydrate structures present in the control glycoproteins selected. *(a)* *N*-linked biantennary complex-type glycan found in fibrinogen and fetuin; *(b)* *N*-linked triantennary complex-type glycan of fetuin; *(c)* *N*-linked high-mannose-type glycan of ribonuclease B; *(d)* *O*-linked hexasaccharide and *(e)* *O*-linked trisaccharide of fetuin.

tested. The proteins were dissolved in 150 mM NaCl and their absorbance at 280 nm was measured by UV spectrophotometry (Cary 1E UV/Vis spectrometer, Agilent Technologies). Concentrations were calculated using an extinction coefficient $E_{1cm}^{1\%}$ = 4.94 for FET and AFET, 15.1 for Fg and 7.1 for RNase A and B, as found in the literature. Serial dilutions of 1, 0.3, 0.1 and 0.03 mg/ml were prepared in printing buffer as for printing needs (sections 3.5.2.2 and 3.5.2.3).

For immobilisation of the high molecular weight polysaccharide mannan, neoglycopeptides were first generated by glutaraldehyde-mediated cross-linking of mannan and the polypeptide poly-L-lysine, responding to the double need of increasing printing efficiency and covalently capturing the probe on functionalised array surfaces.

Ten mg of mannan from *Saccharomyces cerevisae* (Sigma-Aldrich) was dissolved in 1 ml of 50 mM $NaHCO_3$ buffer, pH 9.6, in a glass tube. To favour cross-linking of the glycans, 40 µl of glutaraldehyde solution (25% in water, Grade I, Sigma) was added (1% final concentration) and the mixture was incubated at room temperature for 60 min

with periodic agitation. Subsequently, 136 µl of poly-L-lysine hydrobromide (Mw 30,000-70,000, Sigma-Aldrich) at 1.67 mg/ml in $NaHCO_3$ buffer was added (200 µg/ml final concentration), and the mixture was further incubated for 30 min at room temperature with periodic agitation. Aldehyde groups form a rather stable but reversible bond with Schiff bases (amines). This cross-linking bond was rendered irreversible by addition of $NaBH_4$ traces. Neoglycopeptide preparations were then exhaustively dialysed against milli-Q® water and finally PBS. The hexose content in each preparation was determined by the phenol–sulphuric acid method, as described in section 3.10.2. Dilutions of 6, 3, 2, 1, 0.3 and 0.1 mg/ml of mannan and mannan-poly-L-lysine conjugate were prepared in printing buffer.

Different LPSs were also printed in the arrays, including *Salmonella enterica* serotype minnesota smooth and rough (R345, R_b) LPS (Seydel *et al.*, 1993), purchased from Sigma-Aldrich and Enzo Life Sciences respectively, and O-antigen containing LPSs from *Klebsiella pneumoniae* strain Kp52145, provided by Dr. Enrique Llobet. The LOS from different NTHi strains was isolated as described in section 3.2 and included in the arrays upon quantitation (see sections 3.11.3 and 3.12.2).

To enable probe detection in the array, 1 µl per ml of Cy3® mono-reactive NHS ester (Fig. 3.5, GE Healthcare) was added to all the above described samples. Under neutral pH conditions, the samples are not expected to be covalently labelled, as NHS conjugation to amine groups occurs at pH above 8.5. The dye can be easily naked-eye detected in solution (pink colour), thanks to its high extinction coefficient. Since its excitation maximum is 548 nm and emission maximum is 568 nm, it can be excited by either using the blue (488 nm) or the green (532 nm) laser. Scanning directly upon printing in the arrays allows probe tracing.

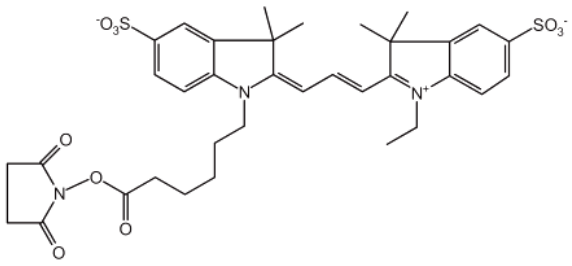


Fig. 3.5 Cy3® mono-reactive dye structure.

3.5.2 Preparation of manual and robotic microarrays

3.5.2.1 Slide surfaces

Several slide surfaces are currently commercially available. Selection of the appropriate surface for the experiment to be performed should be based on probe immobilisation efficiency, spot morphology and array reproducibility. In addition, the slide coating should be of low fluorescence, so that background "noise" is minimised. In this Thesis, nitrocellulose- (NC) and hydrogel-coated slides were used as probe immobilisation surfaces.

Fig. 3.6 16-pad NC-coated glass slide. Adapted from *www.scienion.com*

NC-coating of glass slides (Fig. 3.6) constitutes the most common surface for non-covalent attachment. The immobilisation mechanism is presumably based on the hydrophobic adsorption of the arrayed molecules on the membrane. Compared to non-coated glass slides, the nitrocellulose membrane layer allows a 3-D presentation of the probes and more even morphology of spots. NC slides have been used to array a huge variety of probes, including glycoproteins, neoglycolipids, underivatised or conjugated polysaccharides (Feizi *et al.,* 2003; Wang *et al.,* 2002), up to intact cells (Campanero-Rhodes *et al.*, 2015). A drawback of the NC-coated slides is the strong intrinsic fluorescence of the membrane at wavelengths used for the detection of common green fluorescent dye-signals (e.g. fluorescein or AlexaFluor®488), what limits the detection of immobilised probes or interacting targets.

3D-hydrogel-coated slides (NEXTERION® Slide H, Schott AG) are ideally suited for the covalent attachment of amine-containing probes, including peptides and proteins,

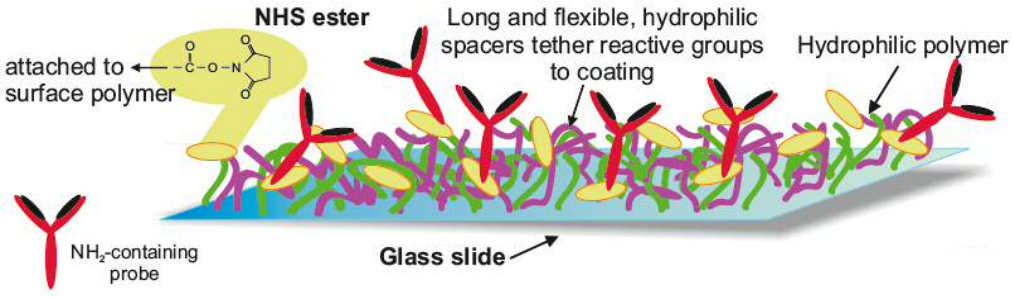


Fig. 3.7 Covalent probe immobilisation on a Slide H surface by amine-reactive chemistry. The protein native structure and biological function is preserved. Inspired from *www.schott.com*.

amino-modified oligonucleotides and amino-linked glycan structures (Alvarez and Blixt, 2006; Blixt *et al.*, 2004; Song *et al.*, 2009). Slide H coating consists of a cross-linked multi-component hydrophilic polymer layer with long and flexible spacers, activated with NHS esters (Fig. 3.7). This permeable polymer coating helps to preserve the native three-dimensional structure of complex biomolecules. Furthermore, slides H produce low background noise, as a low-fluorescence glass is coated with a low-fluorescence polymer.

3.5.2.2 Manual printing

Manual arrays were prepared by printing the probes on single-pad NC-coated glass slides (FAST®-Slide, Whatman) using a manual glass-slide indexing system (VP 470, V&P Scientific) and a floating pin replicator. The replicator contains 4 rows of 8 pins (32 pins in total) on 4.5 mm centres (VP 478A, V&P Scientific) and is ideal for sampling from 384-well microplates (Fig. 3.8). Two slides can be printed at the same time, facilitating the production of duplicates and making printing time effective. The replicator pins (FP1S6, V&P Scientific) have a 6 nl slot. The probe solution is transferred in the slot along with a hanging drop at the tip of the pin and on the sides of the pin (up to ~30 nl in total). Each probe was printed in a four-level dose-response format, in quadruplicate spots, by applying approximately 8 nl per spot. The precise amount finally deposited on the slide depends on the surface tension of both the solution and the slide.

Fig. 3.8 The VP 470 manual glass-slide indexing system. Illustration of the VP 478A floating pin replicator.

3.5.2.3 Non-contact robotic printing

Contact printing microarrayers, first developed in 1995, usually have solid, split or microspotting pins that come into direct contact with the slide surface. One of the main disadvantages of this printing technique is that the slide surface may get damaged during the printing process. To avoid this problem, non-contact robotic arrayers gained ground lately, being characterised by a high-yield performance and great versatility.

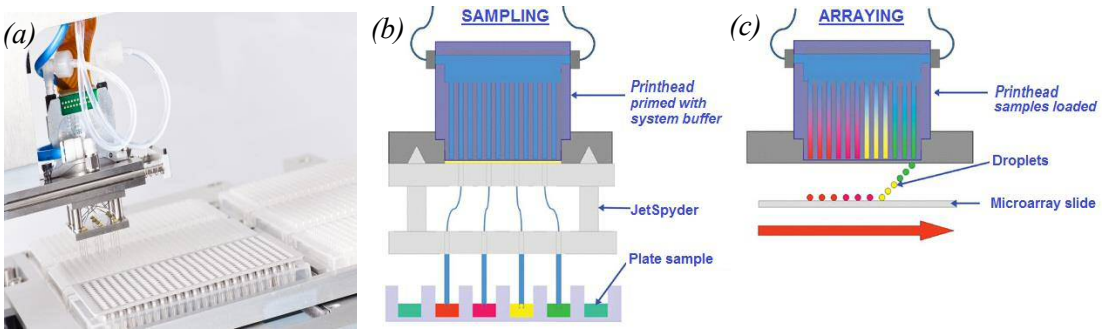


Fig. 3.9 *(a)* Sample loading by JetSpyder into the printhead; *(b)* sampling and *(c)* arraying scheme. Adapted from www.arrayjet.co.uk.

Robotic arrays were printed using a 16-pad layout on either NC- (FAST®-Slide, Whatman) or 3D-hydrogel-film-coated slides, using a non-contact piezoelectric arrayer (Sprint, Arrayjet Ltd, Fig. 3.9) equipped with an environmental control system (JetMosphere™, Arrayjet Ltd) for maintaining temperature and appropriate relative humidity (~50%) conditions, thus providing a consistent environment for array printing.

The arrayer is fitted with a 12-pin loading device (JetSpyder™, Arrayjet Ltd, Fig. 3.9, A, B) with capacity for enabling the printhead to print up to twenty 16-pad slides at a charge. The samples are aspirated through the loading device into the printhead, which has linearly arranged nozzles. After sample loading, the JetSpyder™ is detached from the printhead and sample droplets are ejected from the printhead nozzles after a piezoelectric impulse (Fig. 3.9, *c*). Careful washing between each sample-set print assures correct loading and dispensation. In this Thesis, up to 6 slides have been successfully printed at a time, following the same indexing pattern. Probes were printed in a multi-level dose-response format at customised concentrations by applying 100 pl per spot under precise XYZ control in different patterns.

One limitation of the robotic arrayer is the obligatory use of a high-density printing buffer, containing 47% glycerol and 0.06% (v/v) Triton X-100, to guarantee uniform distribution of the samples and cleanliness of the instrument. Once printed, NC-coated slides were kept overnight in a dry place to allow glycerol evaporation and then stored at 20 ºC. Hydrogel-coated slides were kept overnight under JetMosphere™-control (relative humidity 50%, temperature 25 ºC) and then stored at -20 ºC in sealed boxes in the presence of desiccant.

3.5.2.4 Microarray quality control

In order to monitor printing quality and immobilisation efficiency, slides were scanned upon printing with a GenePix 4200-Autoloader microarray scanner (Axon, Molecular Devices), using the appropriate excitation laser (blue/488 nm for SYTO® 13 or green/532 nm for Cy3®) and emission-detection range (Fig. 3.10). Fluorescence signals were quantified with the GenePix Pro 6.1 software (Molecular Devices).

Microarray stability was examined by scanning slides after different storage periods. In addition, retention of the probes in the arrays after the binding assays was confirmed using the appropriate scanning parameters.

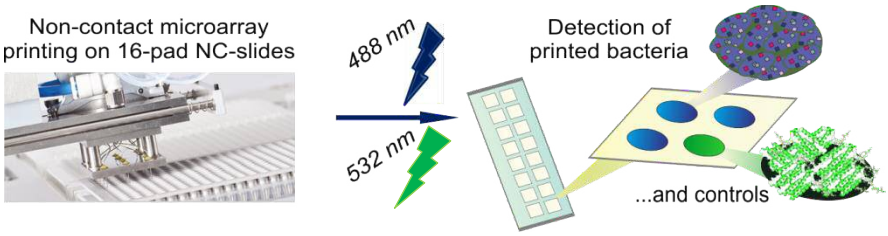


Fig. 3.10 Probes monitoring/localisation upon printing, and microarray quality assessment.

3.5.3 Preparation and biotin labelling of target proteins

Surfactant pulmonary-associated protein D (SP-D) is a member of the collectin family (C-type animal lectins) consisting of four discrete structural domains: a short N-terminal domain, a relatively long collagenous tail domain, a coiled-coil neck domain, and a C-terminal CRD (Crouch, 2000). In this Thesis, N-terminally histidine-tagged human recombinant SP-D (BioVision Inc), consisting of the CRD+Neck domain (18.9 kDa, aminoacids 224-375 of the full-length SP-D sequence), was used without any further modification (Fig. 3.11, *a*).

Human Siglec-14 is a member of the Ig-like family (I-type animal lectins) which are transmembrane proteins with an extracellular N-terminal Ig-like V-set domain followed by a varying number of Ig-like C2-set domains (Crocker and Varki, 2001). For binding assays in solution, a chimeric protein consisting of the human Siglec-14 V-set domain (aminoacids 18-358 of the full-length Siglec-14 sequence) linked by peptide IEGRMD to human IgG1-Fc domain (aminoacids 100-330), expressed in mouse myeloma NS0-derived cell line was purchased from R&D Systems Inc and used without any further modification (Fig. 3.11, *b*).

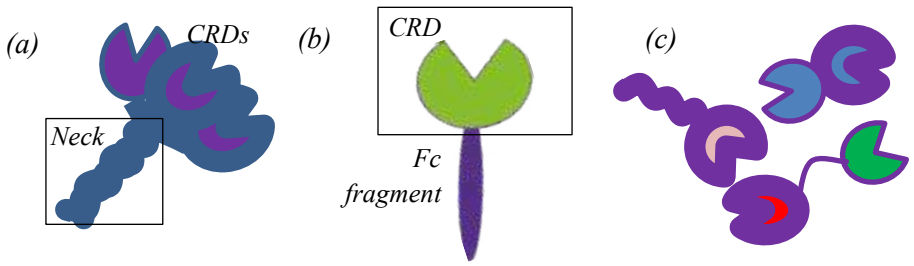


Fig. 3.11 Endogenous lectins used in this Thesis.
(a) Human recombinant SP-D (CRD + neck region); *(b)* human, chimeric Fc-Siglec 14;
(c) homodimeric, chimeric and tandem-repeat human galectins (also see section 4.1.10)

Unless otherwise stated, recombinant galectins used (Fig. 3.11, *c*) were expressed in *E. coli* and purified by affinity chromatography on lactose-Sepharose 4B columns (Kaltner *et al.*, 2008) at the laboratory of Prof. Hans-Joachim Gabius (Lehrstuhl für Physiologische Chemie, Ludwig-Maximilians-Universität, Munich, Germany). For biotin labelling, galectins were exhaustively dialysed at 4 ºC against PBS containing 4 mM β-mercaptoethanol (PBS$_\beta$), followed by a single dialysis against PBS$_\beta$ containing 20 mM lactose, in order to prevent a potential modification upon labelling of free amino groups of the carbohydrate-binding site. Galectins were then incubated with biotinamido-caproate ester derivative (GE Healthcare) in 36 mM $NaHCO_3$ buffer, pH 8.6, at 20 ºC for 1 h under constant shaking, according to the manufacturer's recommendations. After labelling, galectins were exhaustively dialysed against PBS$_\beta$ at 4 ºC and their concentration was estimated spectrophotometrically at 280 nm by using the extinction coefficient calculated from the amino acid sequence with the ProtParam tool (http://web.expasy.org/protparam). The purity of the preparations was assessed by SDS-PAGE, as described in section 3.11.1. Commercial (ATGen Ltd.) recombinant galectins used included N-terminally His-tagged human galectins-3, -4 and -8.

Concanavalin A (jack bean lectin, ConA), *Galanthus nivalis* (snowdrop) agglutinin (GNA), *Ricinus communis* agglutinin (RCA) and *Triticum vulgaris* (wheat germ) agglutinin (WGA) were purchased from VECTOR Labs, and used either unlabelled or biotin-labelled depending on the experiment. For biotin labelling, a protocol similar to that described above was followed, except that no reducing agent was used in the dialysis buffer, and 20 mM α-Me-Man, lactose (Lac) or *N*-acetylglucosamine (GlcNAc) was used to protect the binding site according to the lectin binding specificity (please see Table 3.2).

Commercial biotinylated lectins used were *Arachis hypogaea* (peanut) agglutinin (PNA), *Datura stramonium* lectin (DSL), *Hippeastrum hybrid* lectin (HHL), *Maackia amurensis* lectin-I (MAL-I), *Sambucus nigra* (elderberry) agglutinin (SNA) and *Wisteria floribunda* agglutinin (WFA) from VECTOR Labs; *Griffonia simplicifolia* lectin-I and isolectin B4 (GSL-I, GSL-IB_4), *Helix pomatia* (edible snail) agglutinin (HPA) and *Lycopersicon esculentum* lectin (LEL) from Sigma-Aldrich; *Glycine max* (soybean) agglutinin (SBA), *Limulus Polyphemus* (horseshoe crab) agglutinin (LPA), *Pisum sativum* (garden pea) agglutinin (PSA), *Polyporus squamosus* (mushroom) lectin (PSqL) and *Viscum album* (mistletoe) agglutinin (VAA) from EY Labs.

Rabbit anti-NTHi serum, raised against a mixture of acetone-killed NTHi strains 375, 398 and 2019 as previously described (Morey *et al.*, 2011), was a generous gift of Dr. Junkal Garmendia. Fluorescein-labelled Annexin V, mouse monoclonal anti-polyhistidine (His-tag), clone HIS-1, and anti-human-CD63, clone MEM-259, antibodies were purchased from Sigma-Aldrich. Mouse polyclonal anti-*Salmonella typhimurium* and goat polyclonal anti-*E.coli* O157-lipid A antibodies were purchased from Abcam plc, and biotinylated mouse anti-human IgG1 Fc was from Life Technologies Inc. Biotinylated goat anti-rabbit IgG and goat anti-mouse IgG (Sigma-Aldrich) and rabbit anti-goat IgG (Abcam plc) were used as secondary antibodies. All antibodies mentioned, as well as the anti-NTHi serum, were diluted in the appropriate buffer directly before use without any further preparation.

3.5.4 Microarray binding and competition assays

Specific set ups for the different binding and competition assays, according to the slide coating/functionalization and nature of immobilised probes and/or targets were established.

3.5.4.1 Antibody binding to NTHi and NTHi LOS microarrays

Serial dilutions of SYTO® 13-labelled NTHi clinical isolates and NTHi375 wild type and mutant strains presenting a sequentially truncated LOS (please see section 4.1.1 for description), as well as isolated LOSs and lipid A, were manually or robotically printed in NC-coated glass slides as described in sections 3.5.2.2-3.

NC-pads were hydrated using milliQ® water and directly blocked for 60 min with blocking buffer composed of PBS, 2% bovine serum albumin (BSA) and 0.25%

Tween®20 (both from Sigma-Aldrich). Slides were carefully washed with PBS containing 0.05% Tween® 20 (washing buffer) and subsequently incubated for 60 min with a 1:500 dilution of rabbit anti-NTHi serum, mouse polyclonal anti-*Salmonella typhimurium* antibody or goat polyclonal anti-lipid A antibody in PBS containing 1% BSA and 0.1% Tween® 20 (overlay buffer). After three short washes with washing buffer, the slides were incubated for 45 min with a 1:1000 dilution of the respective biotinylated secondary antibody in overlay buffer. Following three additional washes, the slide was finally incubated for 35 min with streptavidin-AlexaFluor®647 (AF647) conjugate (Molecular Probes™, Life Technologies) at 1 μg/ml in overlay buffer. Finally, slides were thoroughly washed with washing buffer, PBS and then milliQ® water and dried under a nitrogen stream. All incubation steps took place in a dry, dark place, at 20 °C.

Following the binding assay, microarrays were scanned for SYTO® 13 signals of printed bacteria and for AF647 signals, using the GenePix® 4200-AL microarray scanner. AF647 has excitation/emission maxima at 650/668 nm, being therefore excited by a red laser (635 nm). Image analysis and quantification of fluorescence signals was performed with the GenePix® Pro 6.1 software (Molecular Devices) and statistical analysis was carried out with the assistance of Microsoft® Excel.

3.5.4.2 Lectin binding and competition assays

3.5.4.2.1 Binding and competition assays on nitrocellulose-coated slides

Probes including bacteria, LPS/LOS, extracellular vesicles and control (glyco)proteins were printed on NC-coated slides as described in section 3.5.2.3. NC-pads were hydrated using milliQ® water and blocked for 60 min with blocking buffer composed of saline buffer and 0.25% Tween® 20. When working with lectins as targets, BSA was not included in the blocking buffer because it increases the background noise and thus reduces the sensitivity. After blocking, the slides were carefully washed with washing buffer containing 0.05% Tween® 20, and incubated for 75 min with the biotin-labelled lectin of choice in buffer containing 0.1% Tween® 20 (Fig. 3.12, *a*). The buffer was selected according to the lectin preferences as for pH, presence of Ca^{2+} and ionic strength (the default NaCl concentration was 150 mM). PBS, TBS or HBS (10 mM HEPES, pH 7.4, containing 150 mM NaCl) buffer were used as detailed in Tables 3.2 and 3.3. Lectin dilutions were prepared immediately before the assay in the

Table 3.2 Plant and animal lectins used for glycophenotyping in this Thesis

Lectin	Abbreviation	Nominal specificity	Experimental conditions				
			Concentration (μg/ml)	[Binding sites] (μM)[a]	Buffer	Ca^{2+} (10 mM)	Inhibitory hapten
Griffonia simplicifolia lectin I, isolectin B4	GSL-IB4	Gal*α*	171	6	TBS	+	0.1 M Gal
Viscum album agglutinin	VAA	Gal > GalNAc	115	2	PBS	-	0.1 M Gal
Ricinus communis agglutinin	RCA	Gal*β*, Gal*β*(1,4)Glc/GlcNAc	12	0.2	PBS	-	0.1 M Lac
Maackia amurensis lectin I	MAL-I	[Neu5Ac*α*(2,3)]Gal*β*(1,4)GlcNAc/Glc	130	3.5	PBS	-	0.1 M Lac
Glycine max agglutinin	SBA	Terminal GalNAc > Gal	82.5	2.75	PBS	-	0.1 M GalNAc
Griffonia simplicifolia lectin I	GSL-I	Gal*α*, GalNAc*α*	57	2	TBS	+	0.1 M Gal
Arachis hypogea agglutinin	PNA	Terminal Gal*β*, Gal*β*(1,3)GalNAc	180	6.5	TBS	+	0.1 M Gal
Wisteria floribunda agglutinin	WFA	Terminal GalNAc, GalNAc*α*(1,3/6)Gal	34	1.2	PBS	-	0.1 M GalNAc
Helix pomatia agglutinin	HPA	GalNAc > Gal*α*	395	30	TBS	-	0.1 M Gal
Datura stramonium lectin	DSL	GlcNAc*β*(1,4)GlcNAc oligos, LacNAc	64.5	3	PBS	-	0.1 M GlcNAc
Lycopersicon esculentum lectin	LEL	GlcNAc*β*(1,4)GlcNAc oligos	71	4	PBS	-	0.1 M GlcNAc
Triticum vulgaris agglutinin	WGA	GlcNAc > Neu5Ac	9	0.5	TBS	+	0.1 M GlcNAc
Limulus polyphemus lectin	LPA	Neu5Ac*α*(2,3/2,6)GalNAc > Neu5Ac > GlcNac > PCho	40	4.2	TBS	+	20 mM 3´SL
Polyporus squamosus lectin	PSqL	[Neu5Ac*α*(2,6)]Galβ(1,4)GlcNAc/Glc	14	1	PBS	-	0.1 M Lac
Sambucus nigra agglutinin	SNA	Neu5Ac*α*(2,6)Gal/GalNAc, GalNAc, Lac	70	2	PBS	-	0.1 M Lac
Concanavalin A	ConA	Man*α* > Glc*α*	10	0.4	HBS	+	0.2 M *α*-Me-Man
Galanthus nivalis agglutinin	GNA	Terminal Man*α*(1,3)Man	52	4	PBS	-	0.1 M Man
Hippeastrum hybrid lectin	HHL	Man*α*(1,3/1,6)Man	125	10	PBS	-	0.1 M Man
Pisum sativum agglutinin	PSA	Man*α*, Glc*α*, GlcNAc*α*	24.5	1	TBS	-	0.1 M Man

[a]Molar concentration of binding sites was calculated taking into account the number of sites per subunit; [b]3'SL, 3' sialyl-lactose.

Table 3.3 Endogenous lectins used in this Thesis

Lectin	Abbrev.	Experimental conditions		
		Concentration (μg/ml)	[Binding Sites] (μM)	Buffer
Human galectin-1	hGal-1	30	2	PBS_{β}
Human galectin-2	hGal-2	73	5	PBS_{β}
Human galectin-3 (full length)	hGal-3	26	1	PBS_{β}
Human galectin-3 (CRD)	hGal-3tr	18	1	PBS_{β}
Human galectin-4 (full length)	hGal-4	9-18	0.25-0.5	PBS/PBS_{β}
Human galectin-4 (N-domain)	hGal-4N	85	5	PBS_{β}
Human galectin-4 (C-domain)	hGal-4C	8.5-17	0.5-1	PBS_{β}
Human galectin-8 (full length)	hGal-8S	18	0.5	PBS_{β}
Human galectin-8 (N-domain)	hGal-8N	9	0.5	PBS_{β}
Human galectin-8 (C-domain)	hGal-8C	18	1	PBS_{β}
Human galectin-9 (full length)	hGal-9	20	0.5	PBS_{β}
Human Siglec-14-Fc chimera	Siglec-14	1	0.015	PBS
Human surfactant pulmonary-associated protein D	SP-D	6	0.3	HBSC

absence/presence of different concentrations of inhibitors (Table 3.2, see also Chapter 4 of this Thesis). In the case of His-tagged lectins, before incubation with streptavidin the slides were incubated for 1 h with a dilution in overlay buffer of mouse anti-His-tag antibody (1:1,000) and biotinylated anti-mouse IgG secondary antibody (1:300), pre-complexed for 1 h (Fig. 3.12, *b*). In the case of human Fc-chimeric Siglec-14, 1 μg/ml was pre-complexed for 2 h with a 30 μg/ml solution of biotinylated mouse anti-human IgG1 Fc antibody in PBS containing 0.1% Tween® 20 (Fig. 3.12, *c*). The slides were then briefly washed with washing buffer and incubated for 35 min with 1 μg/ml AF647-

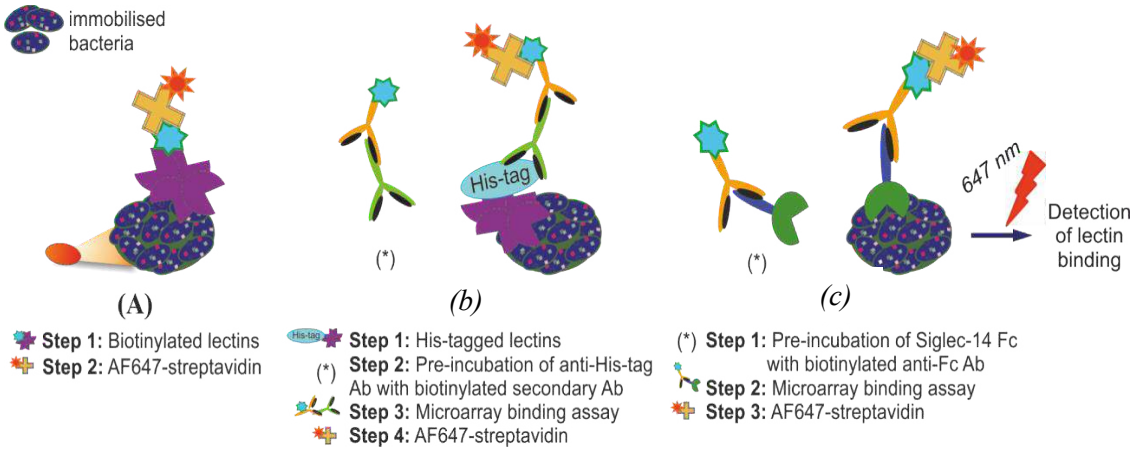


Fig. 3.12 Steps for detecting binding on NC-based microarray experiments. *(a)* Biotin-labelled lectins, *(b)* His-tagged lectins, including incubation with pre-complexed anti-His-tag & biotin-anti-mouse IgG, and *(c)* chimeric Siglec-14-Fc pre-complexed with biotinylated mouse anti-human Fc. Adapted from (Kalograiaki *et al.*, 2016).

streptavidin in overlay buffer for detection of bound biotinylated lectin or antibody.Slides were finally washed thoroughly with washing buffer, Tween® 20-free buffer andthen milliQ® water, and dried under nitrogen stream. All incubation steps took place in a dry, dark place, at 20 ºC. Arrays were scanned with the blue and red lasers for detection of SYTO® 13 and AF647 signals, respectively (Fig. 3.10 and 3.12).

3.5.4.2.2 Binding and competition assays on hydrogel-coated slides

This protocol was followed for assessment of the inhibitory potential of monosaccharides, polysaccharides as well as synthetic glyconanoparticles (section 3.4) and glycofullerenes (for description please see section 4.3.2.2).

Following printing of (glyco)proteins and neoglycopeptides on Slides H, these were washed with milli-Q® water, then with HBS buffer, and any remaining active site was finally de-activated by incubation for 90 min with 50 mM ethanolamine in 0.1 M boric acid, pH 8.5. The de-activated slides were thoroughly washed with milli-Q® water and were ready to use for binding assays. Alternatively they were dried under nitrogen stream and stored at -20 °C, remaining stable for at least 12 months (also see 4.3.1).

Slides were carefully washed with HBS containing 0.1% Tween® 20 (HBS-T) and incubated for 75 min in the same buffer with the biotinylated lectins in the absence or presence of diverse inhibitor concentrations. Hybridised slides were then washed twice with HBS containing 0.05% Tween® 20 and incubated for 35 min with 1 μg/ml streptavidin-AF647 in HBS-T. After washing with HBS-T and milli-Q® water, the slides were dried under nitrogen stream and scanned for AF647 signals.

Alternatively, the de-activated slides were washed and incubated with non-labelled lectins in the absence of inhibitors, using the same protocol. After washing, slides were incubated with 0.1 mM biotinylated heptyl-mannoside (MAF285, Fig. 3.13), in HBS for 60 min at 20 ºC, in the absence or presence of different concentrations of inhibitors.

Detection of lectin-bound MAF285 by incubation with streptavidin-AF647 was performed as described above. Inhibitory activity (IC_{50} values) was determined by semi-

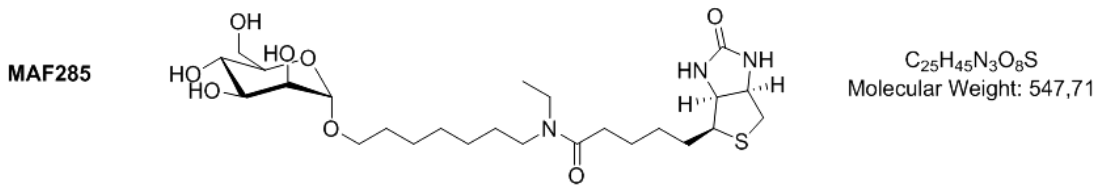


Fig. 3.13 Biotinylated heptyl-mannoside**.**

logarithmic plotting of the percentage of bound MAF285 versus inhibitor concentration, considering as 100% the binding in the absence of inhibitors, and sigmoidal fit to experimental data using the Origin 7.0 software (OriginLab). MAF285 was synthetised in the laboratory of Prof. Stéphane Vincent (Unité de Chimie Organique, Université de Namur-FUNDP, Belgium).

3.5.4.3 Bacteria binding and competition assays

Different mannose-containing glycoproteins and neoglycopeptides were printed in a 16-pad format on hydrogel-coated Slides H (NEXTERION®), followed by slide washing and deactivation as described in the previous section.

E. coli strain UTI89 (*E. coli*) was grown and stained with SYTO® 62 as described in section 3.1. Slides H were incubated for 1 h with 100 µl of SYTO® 62-stained bacteria (red) at OD_{600} = 1, in the absence or presence of different concentrations of natural or synthetic potential inhibitors. After a short wash in HBS and another in milliQ®, the slides were dried under nitrogen stream and scanned with a 635 nm laser for detecting SYTO® 62 signals (Fig. 3.14). The inhibitory activity (IC_{50} values) for each compound was determined by semi-logarithmic plotting of the percentage of bound bacteria versus inhibitor concentration, taking as 100% the binding in the absence of inhibitors, and sigmoidal fit to experimental data using the Origin 7.0 software.

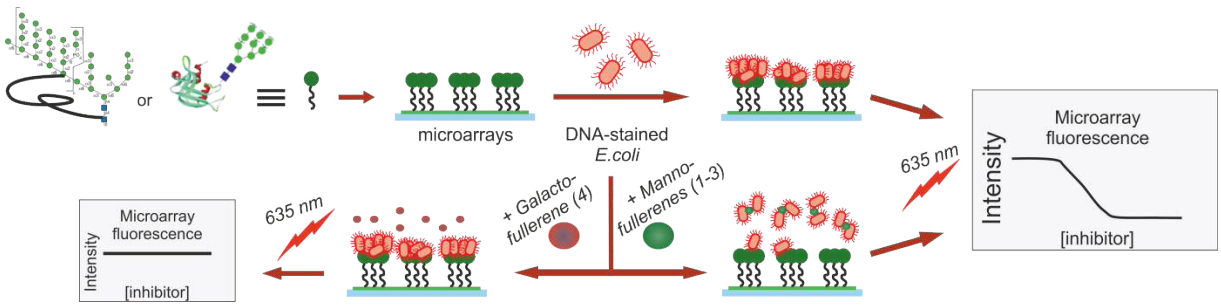


Fig. 3.14 Concept of the bacteria binding and glycofullerene competition experiments on hydrogel slides.

3.6 Piezoelectric and optical biosensor assays

In this Thesis, the term biosensor is applied to sensors using biomolecules as recognition elements that selectively interact with an analyte. A suitable transducer converts the biological response into a digital electronic signal directly proportional to the quantity of bound analyte (Lowe, 1985). Affinity biosensors are characterised by the presence of optical, magnetic or piezoelectric transducers that interconvert light

intensity, mechanical deformation or voltage to measure mass or viscoelastic effects (Turner, 2000). In this work, assays were conducted using either piezoelectric/acoustic (Quartz Crystal Microbalance, QCM) or optical (Surface Plasmon Resonance, SPR) label-free biosensor devices, which permit kinetic evaluation of biological interactions.

3.6.1 Quartz Crystal Microgravimetry

Acoustic sensors comprise one or more vibrating elements creating acoustic waves that propagate on the surface or in the bulk of the resonator (bulk acoustic wave, BAW). The properties (amplitude or frequency) of these waves change when molecules bind to the sensor surface (Marx, 2003).

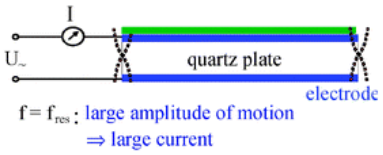


Fig. 3.15 Schematic diagram of a QCM Thickness-Shear-Mode resonator, excited by two electrodes. Adapted from (Johannsmann, 2008)

In particular, quartz crystals attained significance as analytical tool after the discovery that their resonant frequency (f) decreases linearly when additional mass is attached to the sensor (Sauerbrey, 1959). The QCM is a BAW device, consisting of a piezoelectric quartz crystal that resonates when electrically excited by two electrodes (Fig. 3.15) (Janshoff *et al.*, 2000). The electrodes can be one of a number of metals (frequently gold) sputtered on the upper and lower quartz surfaces. By applying an alternating electric field, a mechanical oscillation of characteristic frequency, f, is produced in the crystal. Typically, the continuous resonance mode is utilised and relative shifts in crystal f are measured.

For resonators operating in the Thickness-Shear-Mode (TSM), AT-cut crystals (angle of 35.25º to the z-axis, Fig. 3.16) are commonly used. This geometry provides a stable oscillation with almost no temperature fluctuation in f between 0 and 50 ºC (frequency stability $\Delta f / f \approx 10^{-8}$) (Bottom, 1982). TSM oscillation is primarily lateral to the surface (Fig. 3.16), the amplitude of a vibrating crystal being 1-2 nm (Marx, 2003). Any mass bound to the surface will cause the crystal oscillation frequency to decrease.

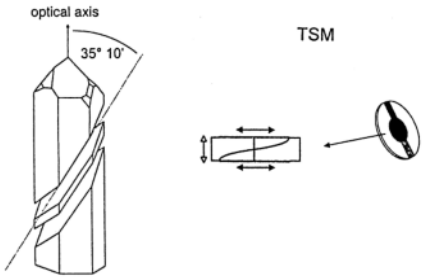


Fig. 3.16 AT-cut of a quartz crystal and an end on crystal view of the thickness-shear-mode of oscillation. Adapted from (Janshoff *et al.*, 2000)

In this Thesis, QCM measurements were performed at Attana AB (Stockholm, Sweden) using a two-channel Attana Cell™ 200 instrument (Fig. 3.17, *a*). The instrument has an

integrated injecting robot (C-Fast) and a peristaltic pump that ensures a continuous flow of running buffer. The flow runs through a pulse dampener to reduce the fluctuations in pressure generated by the pump. The whole system is controlled by the Attester software (Attana AB), which also provides frequency shift information in real-time.

One chip is used per channel, each housing a quartz crystal in a small cylindrical chamber (height: 50 µm, diameter: 5.6 µm, volume: 1.25 µl). For preparation of LNB

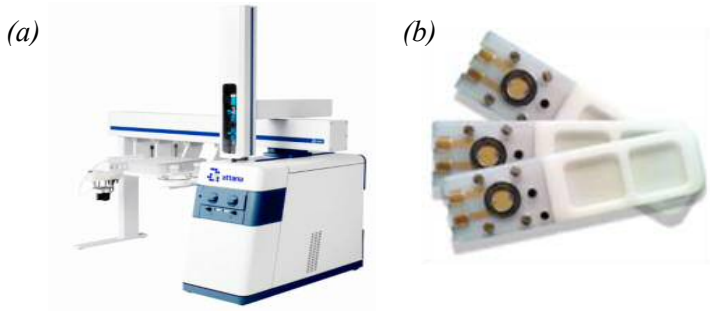


Fig. 3.17 *(a)* Attana Cell™ A200 instrument, C-Fast robot-assisted quartz crystal microgravimetric system; *(b)* LNB-Carboxy Chips.

(Low Non-specific Binding)-carboxyl QCM chips (Attana AB, Fig. 3.17, *b*), quartz crystals were pre-treated as follows. AT-cut quartz crystals of 10 MHz nominal frequency and ~8 mm diameter, sputtered with gold electrodes of 4.5 mm diameter and 140 nm thickness on each side, were cleaned in fresh "piranha" solution (30% H_2O_2/H_2SO_4, 1:3 v/v) for one min, rinsed with milliQ water and washed thrice in $CHCl_3$.

The gold-coated crystal surface was then functionalised with a self-assembled monolayer (SAM) bearing carboxyl groups, by immersion in an alkanethiol derivative (eg. 10-carboxy-1-decanethiol) ethanol solution of 3-5 mM (Bain *et al.*, 1989) and soaking for ~3 h. The method takes advantage of the rapid and strong chemisorption of long-chain alkanethiols on gold surfaces, leading to the formation of an ordered and oriented monolayer (Fig. 3.18). When the monolayer formation was complete, the chip core was washed with ethanol and distilled water, then dried under nitrogen stream. The carboxy-functionalised gold-coated crystal was finally assembled into the plastic chip

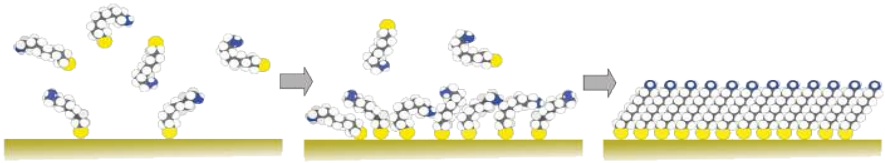


Fig. 3.18 Formation of alkanethiol-based ordered monolayers.
Alkanethiols begin to chemisorb to the gold-coated glass plate (Au-S) when the latter is immersed in an alkanethiol derivative solution. Increase in bond number results in stronger intermolecular interactions, SAMs of higher density and orientation being formed.

holder (Fig. 3.17, *b*).

Two assembled chips were washed with 20 µl milli-Q® water and docked into the Attana Cell™ 200 instrument in channels A and B. Upon chip insertion, the crystals were exposed to the flow of running buffer on a single (gold) surface, avoiding direct contact between liquid and quartz. The temperature was set at 22 ºC and the buffer flow rate at 50-100 µl/min, until stabilisation of the baseline (frequency shift ≤ 0.2 Hz over 600 s). Flow rate was then reduced to 10 µl/min and a fresh activation mix containing 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) and 0.05 M sulfo-NHS (Amine Coupling Kit, Attana AB) was injected for 5 min (Fig. 3.19, *a*). Flow was switched again to running buffer (PBS or HBS containing 0.005% Tween® 20). Lectins to be covalently coupled (ConA and GNA), were diluted at ~50 and 100 µg/ml, respectively, in 10 mM acetate buffer, pH 4, below their isoelectric point (6 for ConA and 4.5 for GNA) in order to facilitate NHS-assisted amine-coupling, and injected twice over both channel A and B surfaces for a total contact time of 10 min. Activation with EDC/sulfo-NHS results in the replacement of the carboxyl groups by more negatively charged sulfo groups (Fig. 3.19, *b*), enhancing initial electrostatic attraction and resulting in efficient protein coupling.

The complete immobilisation process was monitored by frequency shifts (Fig. 3.19, *a*). Following immobilisation, a de-activation solution (1 M ethanolamine, pH 8.5) was injected in both channels for 5 min, and finally, flow was switched to running buffer at 20 µl/min over the lectin-derivatised surface till stabilisation of the baseline. Coupling of ConA on the chip surfaces resulted in a shift of 60-80 Hz in the frequency of the

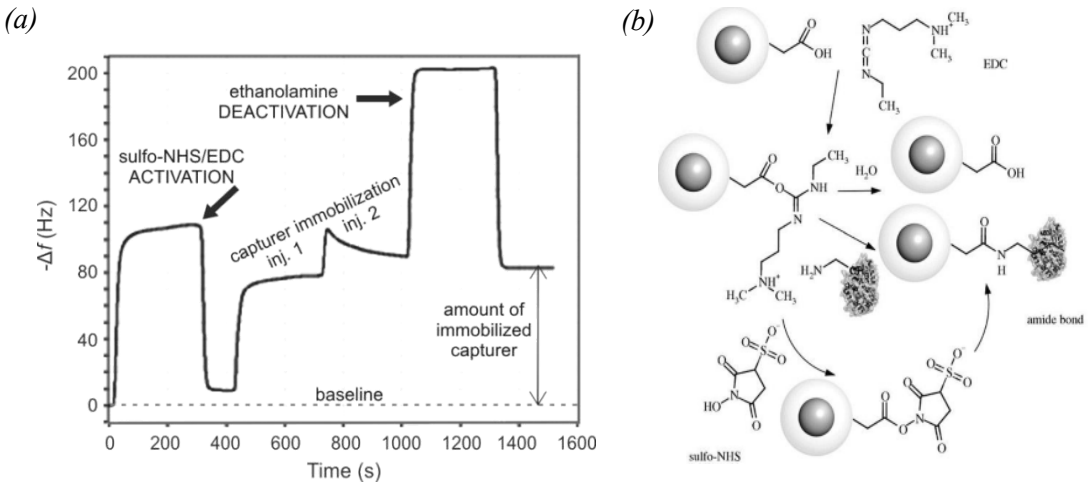


Fig. 3.19 *(a)* QCM sensogram for lectin immobilisation by two consecutive injections in LNB-Carboxy chips with previous EDC/sulfo-NHS activation, as indicated; *(b)* EDC/sulfo-NHS amine coupling chemistry.

quartz crystal resonator versus ~50 Hz for GNA.

Following the Sauerbrey law, the frequency value is decreased with the enhancement of mass on the QCM electrode surface. However, the specific set-up of the Attana instrument includes a third crystal inside the machine which is used as reference, and the reported frequencies correspond to the frequency of this internal crystal (fixed for each instrument) minus the frequency of the experimental crystal. When the mass is increased at the experimental crystal, its frequency is reduced, so after subtraction from the frequency of the internal crystal the final observed frequency shift yields positive values. The read-out from the detector is thus given as -Δ*f* values (Fig. 3.19*a*, and Fig. 3.22).

For QCM analysis of lectin binding to EVs, two strategies were followed. First, the interaction of EVs directly injected over immobilised GNA was monitored. To this end, 1:10 dilutions in PBS of THP-1-derived exosome and MV- preparations were injected in duplicate (injection volume 250 µL). An activated/deactivated chip without GNA immobilisation was used as reference in channel B. After each injection, the GNA-derivatised chip was regenerated with glycine 10 mM, pH 2.5.

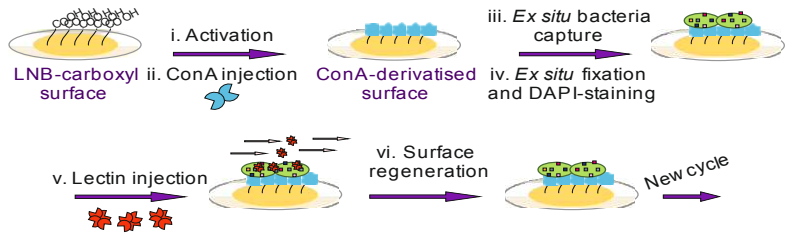


Fig. 3.20 Preparation of bacteria-QCM chips. Adapted from (Kalograiaki *et al.*, 2016)

Second, lectin-captured EV chips were developed for the analysis of lectin binding to the EV surface (sandwich assay). Following GNA derivatisation, as described above, both A and B chips were undocked from the machine and disassembled to permit *ex situ* EV capture. Chip A was then incubated for 1 h with 50 µl of THP-1 exosome stock solution (see section 3.3). Chip B, in this case also GNA-derivatised, was incubated under the same conditions with 50 µl of PBS to be used as reference. Following incubation, the surfaces were washed 3 times with PBS, containing 0.05% Tween-20 (running buffer), fixed for 10 min at 4 ºC with 3.7% (v/v) formaldehyde (Sigma-Aldrich, 37% stock solution diluted in PBS), and washed again extensively with running buffer. The chips were then reassembled, docked into the instrument, and the

flow rate was set at 20 µl/min at 22 ºC, until stabilisation of the baseline (frequency shift ≤ 0.2 Hz over 600 s).

For QCM analysis of lectin binding to bacteria, lectin-captured bacteria chips were developed in a similar manner. ConA was selected as bacteria capturer. Following ConA-derivatisation, both chips were undocked from the machine as described for EVs to permit *ex situ* bacteria capture (Fig. 3.20). Chip A was incubated for 30 min at 22 ºC with 50 µl of a NTHi suspension of OD_{600} = 1, corresponding to ~1×10^9 bacteria/ml, in PBS (cultured and fixed as described in section 3.1). Chip B was incubated under the same conditions with 50 µl of PBS as reference. Chips were washed and fixed with formaldehyde as described above.

Then, the bacteria-derivatised surface was incubated with a 0.1 µg/ml DAPI solution in running buffer for 15 min in the dark. DAPI (4′,6′-diamidino-2-phenylindole, Invitrogen) is a membrane permeable dye, not interfering with membrane molecules, used for staining bacterial nucleic acids. To evaluate the cell coverage of the surfaces, fluorescence microscopy pictures were obtained using an Eclipse 80i microscope (Nikon) equipped with 340-380 nm, 400 nm, and 435-485 nm excitation filter, dichroic mirror and emission filter, respectively. The chips were then reassembled, docked into the instrument and left overnight at 20 µl/min flow rate and 22 ºC, to permit stabilisation of the baseline.

Once the chip frequency was stabilised, full binding kinetic analysis was performed for selected lectins over the fixed bacteria/EVs chips. Increasing concentrations of the lectins in running buffer were injected for 84 s in consecutive duplicates, each injection followed by one of running buffer for 600 s. For information on the concentration range

Table 3.4 Lectins used in QCM kinetic analyses

Lectin	MW (kDa)	Concentration range	
		in µg/ml	in nM*
RCA	120	0.15-5	1-42
VAA	115	0.15-5	1.31-43.8
PNA	110	73-330	660-3,000
WGA	36	6.25-100	175-2,800
ConA	104	10-150	100-1,500
GNA	52	34-312	660-6,000

For lectin specificity please see Table 3.2. *For calculation of the molar concentration, the quaternary structure molecular weight was considered for each lectin.

tested, please see Table 3.4. After each cycle, a pulse injection (30 s) of 10 mM glycine, pH 1.2-1.5 (for bacteria-chips also containing 0.5 M NaCl) led to efficient regeneration, surfaces being immediately re-equilibrated with running buffer (Fig. 3.21). In the competition experiments, the lectins were preincubated with their specific hapten for 1 h in running buffer before injection.

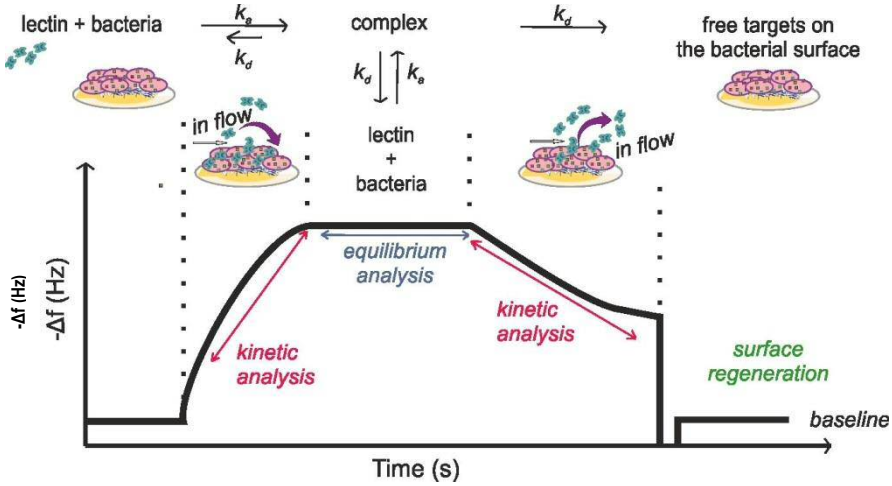


Fig. 3.21 QCM sensogram corresponding to the interactions taking place upon lectin injection over a bacteria-derivatised sensor chip. Regeneration clears analyte bound molecules.

Experimental data were collected using Attester software (Attana AB) and analysed with the evaluation software provided by the manufacturer along with TraceDrawer (Ridgeview). In order to calculate the kinetic parameters (k_{on} and k_{off}) and affinity constant (K_D), the experimental data were fitted using 1:1 or 1:2 models (for details, please see sections 4.1.4 and 4.2.5). Real-time biosensor response for the reference surface (chip B) was subtracted from sensograms obtained for the chip A, ligand surface.

In order to assess chip quality and regeneration capacity, ConA-bacteria chips were removed from the device after the binding analysis and retention of the bacteria was evaluated *ex situ* by DAPI imaging using fluorescence microscopy, as described before. In addition, lectin binding to the bacteria chips was visualized by incubation for 30 min with 150 µg/ml fluorescein-labeled ConA (FITC-ConA) in running buffer. After thorough washing, fluorescence microscopy pictures of the surface were obtained using a Nikon microscope Eclipse 80i equipped with 465-495 nm, 505 nm, and 515-555 nm excitation filter, dichroic mirror and emission filter, respectively.

3.6.2 Surface Plasmon Resonance

The SPR phenomenon occurs when surface plasmon waves (SPWs) related with collective excitations of electrons are excited by light at a metal/liquid interface (Homola, 2003; Kretschmann and Raether, 1968; Otto, 1968; Raether, 1988). Most of the commercially available SPR-sensing devices still use the "attenuated total reflection" geometry proposed by Kretschmann and Raether in the late 60s (Fig. 3.22). This consists in directing plane polarised light from a laser source through a high refractive index (RI) prism, and totally reflect it at the base of the prism where a thin gold film (~50 nm) is deposited (De Crescenzo *et al.*, 2008) (Fig. 3.22).

Total internal reflection (TIR) occurs at an interface between non-absorbing media of different RI (Mirabella, 1985). When a light beam propagating in a medium of high RI meets the interface with a medium of lower RI and is totally reflected, it propagates back into the high RI medium. Although the reflected beam does not lose any net energy across the interface, it leaks electrical field intensity, called evanescent field wave, into the low RI medium. The amplitude of this evanescent field wave decreases exponentially with distance from the interface into the medium. If the TIR-interface is sputtered with a layer of a suitable conducting material (usually gold), the evanescent field wave may penetrate the metal layer and excite electromagnetic SPWs propagating within the gold surface, thus creating an enhanced SPR-evanescent wave. For energy of the photons pumping on the thin gold surface to be transferred to surface plasmons, the tangential *x*-component of the incident optical wave vector ($\boldsymbol{k}_x$) has to match the SPW vector, $\boldsymbol{k}_{sp}$, in magnitude and direction. This resonant absorption of energy, causing plasmon excitation, leads to the appearance of a sharp resonant dip in the reflected light intensity at a definite angle (θ_{spr}) (Fig. 3.22).

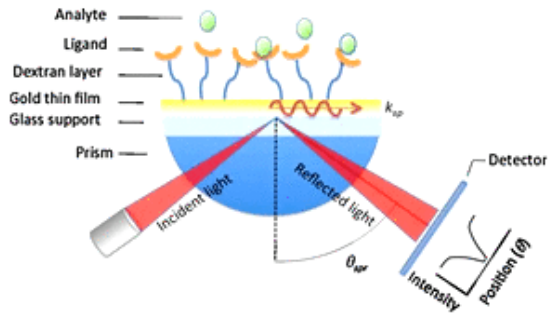


Fig. 3.22 SPR detection principle.
When surface plasmon resonance occurs, a dip in the reflected light is observed at a defined angle (θ_{spr}). Time-dependent accumulation of analyte mass within the vicinity of the surface results in variations of θ_{spr} that are followed in real time. Adapted from (De Crescenzo *et al.*, 2008).

The wave vector of the light, $\boldsymbol{k}_x$, can be adjusted to match that of the surface plasmon by controlling the angle of incident light (Homola, 2008). On the other hand, the SPW vector depends on the RI of the conductor and of the dielectric within the decay length of the wave from the solid–solution interface (typically ~300 nm). Most of the SPR sensors, including Biacore (GE Healthcare), employ this property to monitor binding events in the close vicinity of the gold layer. The binding of an analyte will produce a local increase in the RI and, consequently, the angle of incidence required for resonance (θ_{spr}) is altered. θ_{spr} changes upon binding events are then translated into a resonance signal (Fig. 3.23), where one resonance/response unit (RU) is equal to a θ_{spr} shift of 10^{-4} deg. The displacement in the angle is proportional to the analyte mass and the detection limit achieved may be down to 10^{-7} RI units in terms of the RI shift, corresponding to about 0.1 pg/mm^2 of analyte accumulating on the sensor surface (De Crescenzo *et al.*, 2008).

The real-time response of the SPR experiment is usually presented in the form of a sensogram (Fig. 3.23). At the start of the experiment, the surface has not yet been exposed to analyte molecules and the RU value (baseline) corresponds to the starting critical angle I. When analyte molecules are injected and bind to the surface, the RI and associated θ_{spr} (II) change, leading to an increase in the resonance signal (Fig. 3.23).

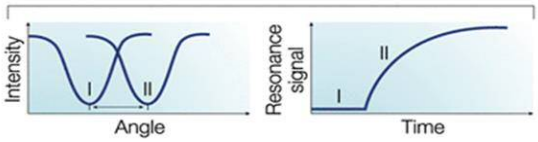


Fig. 3.23 Functional modes of SPR detectors. The reflection angle θ_{spr} is dependent on the RI of the medium within 300 nm of the gold surface (I). Upon an analyte-binding event, the RI is translated into an angle shift (II). The signal is finally represented as resonance units versus time.

In this Thesis, SPR experiments were performed using a Biacore® X-100 biosensing device (GE Healthcare) located at Instituto de Química Médica, Centro de Química Orgánica "Lora Tamayo" (CSIC, Madrid, Spain). In Biacore® instruments, the incident light is focused into a wedge-shaped beam simultaneously providing a continuous interval of $\boldsymbol{k}_x$ vectors that covers the working range for the SPW vector, $\boldsymbol{k}_{spr}$. Changes in the SPR angle are evidenced by changes in the position of the reflected intensity dip (from I to II, Fig. 3.23). The analyte is supplied in a controlled fashion to the sensor surface through a microfluidic system.

Many types of sensor-chips are commercially available, each covered with a different surface matrix. In this work, we used Biacore® CM5 chips, which consist in a glass surface covered by a thin gold layer (~50 nm thick). The latter is first covered with a

covalently bound monolayer of alkanethiol molecules, which is functionalised with a matrix of carboxymethylated dextran, a flexible unbranched carbohydrate polymer forming a thin surface layer, ~100 nm thick (Fig. 3.24). The dextran matrix provides a hydrophilic environment favourable to biomolecular interactions, as well as a tridimensional, well-defined chemical basis for covalent attachment of biomolecules. The sensor chip has two channels: one is a reference channel (untreated) while the second is used for target immobilisation. Kinetic analysis of the binding event can be performed by monitoring the SPR angle as a function of time.

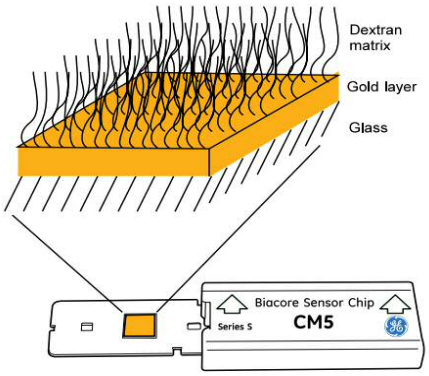


Fig. 3.24 Schematic illustration of the sensor chip surface of the CM-series. Source: Biacore® website

For this study, ConA was immobilised on the chip as follows. The carboxymethylated-dextran surface was first activated for 7 min with a 0.2 M EDC/0.05 M NHS mix injection at a flow rate of 10 µl/min. Then, a 50 µg/ml ConA dilution in 10 mM acetate buffer, pH 4, was injected for 2 min, followed by de-activation with 1 M ethanolamine-HCl, pH 8.5, for 7 min (also serving to remove non-covalently bound ligand), Fig. 3.25. ConA immobilisation resulted in a response of 2700-2900 RU. Upon baseline stabilization, increasing concentrations (0.25-10 µM) of α-D-mannose- as well as galactose-bearing glycofullerenes in HBS were injected for 80 s in both channels at a flow rate of 50 µl/min, at 22 °C. Each fullerene injection was followed by continuous buffer flow for 100 s (flow rate 50 µl/min, dissociation phase), and a single injection of regeneration solution (10 mM glycine-HCl, pH 3.0) for 20 s at 10 µl/min to remove any remaining bound fullerene from the chip surface, then switching again to running buffer at the initial flow rate (50 µl/min), Fig. 3.25.

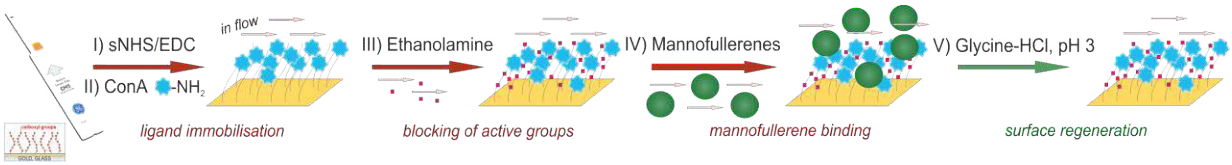


Fig. 3.25 Steps followed for the kinetic analysis of mannofullerene binding to immobilised ConA by SPR.

Response data were collected at real-time, analysed with the Biacore® X-100 evaluation software and plotted as response shift versus analyte concentration. The interaction apparently followed pseudo-first order kinetics, and the dissociation constant (K_D) was calculated in equilibrium following a 1:1 interaction model, using the equation:

$$R_{eq} = \frac{R_{max} \times [A]}{[A] + K_D}$$

where R_{eq} corresponds to the biosensor response in equilibrium (RU), R_{max} to the maximum response registered and [A] to the injected analyte (fullerene) concentration. The bulk index (the difference in RI between the running and sample buffer) was zero in all cases.

3.7 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) measures directly the energetics of a binding process upon mixing of two components, including protein-ligand interactions (Freire *et al.*, 1990; Wiseman *et al.*, 1989).

A power compensation ITC instrument has two identical coin-shaped cells, sample and reference, enclosed in an adiabatic shield (Fig. 3.26). A feedback control system monitors the difference in temperature between the two cells through a semiconductor Peltier sensor device sandwiched between them. In order to maintain the difference in temperature constant, controlled power is continuously supplied to the sample cell, what determines the baseline level.

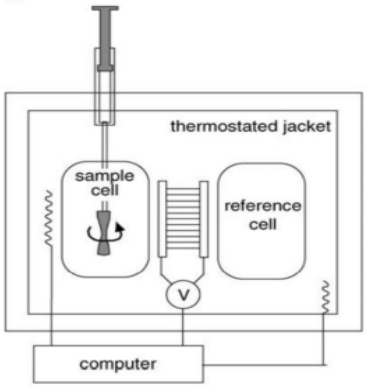


Fig. 3.26 Schematic illustration of a power compensation ITC instrument. Adapted from (Velázquez-Campoy *et al.*, 2004)

One of the reactants (usually the macromolecule, M) is placed in the sample cell and the other one (usually the ligand, L) is added to the cell in a stepwise fashion with the use of a syringe, at the same time stirring the solution in the sample cell to achieve fast mixing (Fig. 3.26). Upon each titrant injection, a certain amount of macromolecule/ligand complex (ML) is formed, depending on the affinity and the concentration of reactants in the sample cell. The reaction may be accompanied by the release (exothermic reaction) or absorption (endothermic) of heat, what causes a difference in temperature between the sample and reference cells. The feedback system then responds by either lowering or rising the thermal power applied to the sample cell, so that the temperature balance is restored and the system reaches equilibrium. Hence, the measured signal is the power applied (μcal/sec) to return to the baseline and shows a typical deflection pattern in the form of a peak (Fig. 3.27, upper panel). Using the baseline as reference, integration of

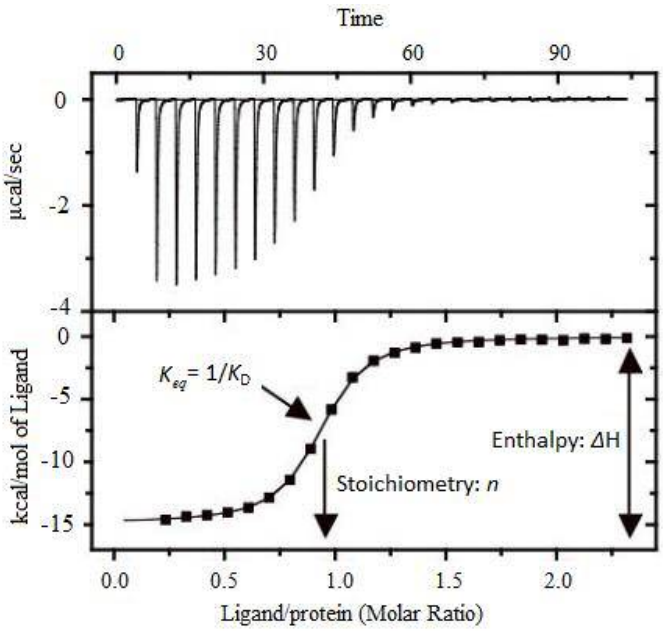


Fig. 3.27 Representative ITC dataset.
In the upper panel each peak represents one injection. In the lower panel, the fit obtained by nonlinear regression analysis of the data can be observed.

the area under the peak provides the amount of heat associated with each injection, which is proportional to the complex formation.

For the binding parameters to be determined, the reaction heat should be both measurable for each injection and also vary for subsequent injections (Fig. 3.27, upper panel). As the degree of saturation of the macromolecule in the cell increases, the heat signal diminishes until only the background heat, due to ligand dilution or mechanical effects, i.e. liquid friction, is observed (Velázquez-Campoy *et al.*, 2004).

Analysis of the heat change per mole of injected ligand as a function of the ligand/protein ratio (Fig. 3.27, lower panel) using the appropriate binding model provides a direct estimation of the binding (equilibrium) constant, K_{eq}, the enthalpy change of binding, ΔH, and the binding stoichiometry, n, whereas the entropy change of the binding, ΔS, can be calculated using the equation:

$$\Delta G = -R \times T \times \ln K_{eq} = \Delta H - T \times \Delta S$$

where R is the universal gas constant (1.987 cal $K^{-1}mol^{-1}$) and T the absolute temperature (K).

ITC analyses of lectin-carbohydrate interactions were carried out either in a MicroCal™ VP-ITC power compensation calorimeter (GE Healthcare) equipped with a 1400 µl sample cell and a 250 µl Hamilton glass syringe, or in a Malvern Microcal PEAQ-ITC instrument equipped with a 200 µl sample cell and a 50 µl Hamilton glass syringe. The lectin of interest was dissolved in HBS, pH 7.4, and exhaustively dialysed against the same buffer. Protein concentration was calculated spectrophotometrically at 280 nm using an extinction coefficient $E_{1cm}^{0.1\%}$ = 1.24 for ConA and 1.80 for GNA.

Lectin solutions were loaded in the sample cell at a monomer concentration of ~300 μM, while the ligands were loaded in the syringe. Ligands included *α*-Me-Man, Man(*α*1,3)man-*α*-OMe (Carbosynth), both at 15 mM, and different glycofullerenes at 0.2 mM (see section 4.3.2.2 for molecules description). A monovalent ligand comprising a *α*-D-mannose moiety and a pegylated linker, used for attachment to the fullerene C_{60} core, was also evaluated (11.7 mM, as calculated by the phenol-sulphuric assay, section 3.11.2). The last buffer used for protein dialysis was employed to prepare the ligand solutions to minimise heat artefacts derived from solvent mixing, and both protein and ligand solutions were degassed prior to filling the cell and the injection syringe to avoid the generation of bubbles and, therefore, spurious heat signals (Freyer and Lewis, 2008). The ITC experiments were run at 22/25 °C and were set to deliver approximately 20 injections at 300-sec intervals, the precise volume series being adjusted according to the experiment. For each protein, one or two titration experiments, depending on ligand availability, were run followed by one ligand dilution experiment, in which the ligand solution was added to dialysis buffer in the cell. The dilution-corrected integrated data were fit with a suitable binding model using the Origin-6™ software (OriginLab) provided with the MicroCal VP-ITC instrument. ConA functional monomer concentration, previously determined by titration with *α*-Me-Man was used as input for ligand concentration in the fitting of dodecamannosylated fullerene titrations, carried out using the "ligand-in-cell" option.

3.8 Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a versatile biophysical technique that permits the elucidation of molecular recognition events at atomic resolution, including the recognition of sugars by their receptors (Weis and Drickamer, 1996). The development of approaches leading to structural determination, as well as the study of the conformation and dynamics of complex oligosaccharides, has contributed to the expansion of its use. Of note, NMR techniques provide the possibility to work in solution and in a vast range of pH, thus in conditions akin to physiological.

All nucleons (neutrons and protons) composing any atomic nucleus, have the intrinsic quantum property of spin. All nuclei are electrically charged (positively, off the protons) and many of them have spin, when the total spin quantum number *S* is not zero. Such nuclei generate a magnetic field of a magnetic moment, μ (Fig. 3.28, *a*), proportional to the spin but also depending on the gyromagnetic ratio γ, characteristic for each nucleus

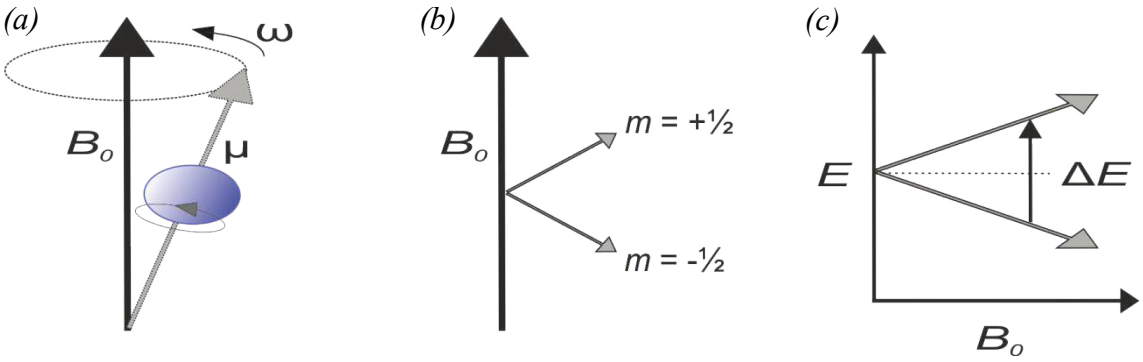


Fig. 3.28 The basis of Nuclear Magnetic Resonance. *(a)* Precession of a $m = +½$ nucleus with an associated magnetic moment, μ, around the external magnetic field vector, B_0, under a Larmor frequency, ω; *(b)* two possible momentum states of a nucleus with $S = ½$ exist when an external magnetic field, B_0, is applied; and *(c)* the dependence of the energy difference (ΔE) between the two momentum states on the external magnetic field strength.

($\mu = \gamma \times S$). The angular momentum corresponding to the nuclear spin is quantised, so that its orientation can only adopt a limited number of values. For a given nucleus of spin S, there exist $2S+1$ possible angular momentum states, of values $-S$ to $+S$. Thus, for nuclei of spin $S = ½$, such as ^{1}H, ^{13}C or ^{15}N, there are two possible spin states: $m = +½$ and $m = -½$ (Fig. 3.28, *b*). In the absence of an external magnetic field, both states have the same energy and the nuclei are equally distributed between them. However, when an external magnetic field (B_0) is applied, this equilibrium is broken and the energy difference between the momentum states is dependent on the external field strength (please see below, *eq. 1*, where $\hbar$ is Planck's constant divided by 2π). As a result, states of lower energy will always have a higher probability of being occupied or, in other words, a slightly larger fraction of the nuclei is present in the lower energy spin state following the Boltzmann distribution (*eq. 2,* where k is the Boltzmann's constant). As it can be appreciated from *eq. 2*, the ratio of nuclei populations at certain temperature only depends on the states' energy difference.

$$\Delta E = \gamma \hbar B_0 \quad \textit{(eq. 1)} \qquad\qquad \frac{N_{low}}{N_{high}} = e^{-\Delta E / kT} \quad \textit{(eq. 2)}$$

The signal in NMR spectroscopy results from the difference between the energy absorbed by the spins that make a transition from the lower to the higher energy state, and the energy emitted by the ones that make a transition from the higher towards the lower energy state. The magnetic moment of each nucleus precesses around the external magnetic field vector, B_0 (Fig. 3.28, *a*). The precession frequency, called Larmor frequency (ω_0), is the frequency of resonance for each nucleus and is equivalent to the energy difference between the two energy levels, ΔE (see below, *eq. 3*). Thus ω_0

depends on both the gyromagnetic ratio and the field strength: the stronger the field, the bigger ΔE and the nucleus precession frequency (Fig. 3.28, *c*).

$$\omega_0 = \gamma \times B_0 \rightarrow \omega_0 = \Delta E/\hbar \quad (eq.\ 3)$$

Nonetheless, local perturbations experienced by every nucleus exist, mainly due to the shielding exerted by nearby electrons. As expected, nuclei with different chemical environments are nonequivalent, precess at slightly distinct frequencies ω (*eq. 4*) and this constitutes the basis of NMR spectroscopy.

$$\omega_0 = \gamma \times (B_0 - B_i) \quad (eq.\ 4)$$

For practical reasons, the use of chemical shifts (δ, expressed in parts per million, ppm) has been established and extensively used instead of Larmor frequencies,

$$\delta = \left(\frac{\omega - \omega_{ref}}{\omega_{ref}}\right) \times 10^6 \quad (eq.\ 5)$$

where ω is a Larmor frequency for a given nucleus, and ω_{ref} is the Larmor frequency of a nucleus in a reference compound, such as tetramethylsilane (TMS).

Nuclei found close to one another, exert an influence on each other's effective magnetic field and this effect shows up in the NMR spectrum when they are nonequivalent. If the distance between nonequivalent nuclei is less than or equal to three bond lengths, this effect, called spin-spin or *J*-coupling, is observable.

3.8.1 Spectral assignment of oligosaccharides

TOtal Correlation SpectroscopY (TOCSY) experiments are bidimensional, homonuclear experiments used to detect spins that are either directly or indirectly coupled to each other through a chain of J-couplings, i.e. belonging to the same spin-system (Bax and Davis, 1985). Practically, cross-peaks of the coupled protons are observed. In the case of oligosaccharides, each sugar residue is an isolated spin-system, and TOCSY experiments would allow for the differentiation and identification of all the protons of a specific sugar residue. Such ability is achieved by inserting a repetitive series of pulses that cause isotropic mixing during the mixing period. Longer mixing times cause a polarisation spread-out through an increasing number of bonds.

For their part, Nuclear Overhauser Enhacement (NOE) and Rotating-frame Overhauser Enhancement (ROE) spectroscopies allow for the establishment of inter-residue

connectivities (Carver, 1991; Woods, 1995). Here, magnetisation is transferred through space in dipole-dipole interactions and the nuclear Overhauser cross-relaxation, observed between spatially-vicinal spins, is used to establish the correlations. In the case of oligosaccharides (MW below 2 kDa), their rotational correlation time falls into a range that turns NOE undetectable. Rotation of the equilibrium magnetisation onto the *x* axis and inhibition of precession by spin-locking enables the establishment of correlations for small molecules (ROESY).

Finally, heteronuclear experiments are useful for achieving the splitting of signals that are overlapping in the ^{1}H dimension, not unusual in carbohydrates due to their poor spectral dispersion (δ=3.5-5 ppm). The most commonly used is the ^{1}H-^{13}C Heteronuclear Single-Quantum Coherence (HSQC), providing a 2D mapping of pairs of directly bound C-H nuclei. NMR spectral elucidation may often be aided by assignments already reported in the literature or stored in databases, such as the Biological Magnetic Resonance Data Bank (BMRB) (Ulrich *et al.*, 2008) or, the computationally simulated, CASPER (Jansson *et al.*, 2006).

For the spectral assignment of the oligosaccharides (OS) derived from NTHi, the isolated LOS of NTHi375Δ*ompP5* or Rd strain was first hydrolysed in 1% acetic acid in milliQ® water (v/v) at 100 °C for 3 h with frequent vortexing. This afforded an insoluble lipid A and core OS fractions in the supernatant. Upon completion of the hydrolysis, an equal volume of milliQ® water was added to the hydrolysate which was then stored for 36 h at 4 °C and later centrifuged at 8,000×g at 4 °C. The recovered supernatant was lyophilized and further used for the NMR experiments. The process was repeated for different LOS preparations.

All NMR spectra were acquired at 310 K (37 °C) on a 600 MHz Bruker Avance spectrometer equipped with a cryoprobe and processed with TopSpin 3.0 software (Bruker). NMR samples were prepared in 99.9% deuterium oxide (D_2O) buffer containing 5 mM phosphate and 200 mM NaCl at pD 7.2 (uncorrected value). The ^{1}H and ^{13}C NMR spectra of the OS were assigned using a combination of TOCSY (*dipsi2phpr*), ^{1}H-^{13}C HSQC (*hsqcedetgp*) and ROESY (*roesyphpr*) of standard pulse sequences, as indicated. TOCSY experiments were performed with 20 and 70 ms mixing times, and the ROESY spin-lock module was of 300 ms. Chemical shifts and coupling constants are reported in section 4.1.6.1. The ^{1}H chemical shifts were

referenced to the residual water signal using the equation δ (ppm) = 5.051 − 0.011 × T^e (°C), (Gottlieb *et al.*, 1997).

3.8.2 Saturation transfer difference

During the last decade, saturation transfer difference (STD) spectroscopy (Mayer and Meyer, 2001) has been extensively used for the evaluation of ligand binding to biomolecules. It is a ligand-based method that involves the comparison of two ^{1}H-NMR spectra of the same sample solution, acquired either with (*on-resonance* spectrum) or without (*off-resonance* spectrum) protein saturation.

The *off-resonance* spectrum, equivalent to a reference ^{1}H-NMR spectrum of the complex, is obtained by irradiating at regions devoid of both ligand and receptor resonances, i.e. at 100 ppm. On the other hand, the *on-resonance* experiment involves selective irradiation of a subset of the protein protons by a train of pulses, directed towards frequencies devoid of ligand signals. Typically the irradiation frequency is either set between -1 and 1 ppm targeting the aliphatic side-chain protons (Ile, Val, Leu), or in the aromatic region (6-8 ppm). The magnetisation propagates via ^{1}H-^{1}H cross-relaxation pathways from the irradiated protons to the rest of the protein, and from the protein to any ligands in direct contact with it. A decrease should be thus observed in the intensity of the saturated proton signals (belonging to both the receptor and the bound ligand; the closest they are to the binding site, the stronger their signal decreases. The rest of signals are not affected. The difference between the *off-resonance* and the *on-resonance* spectra yields a saturation spectrum in which the saturated nuclei signals are observed and allow for ligand identification or epitope-mapping. Application of protein spin-lock or solvent suppression modules is also possible.

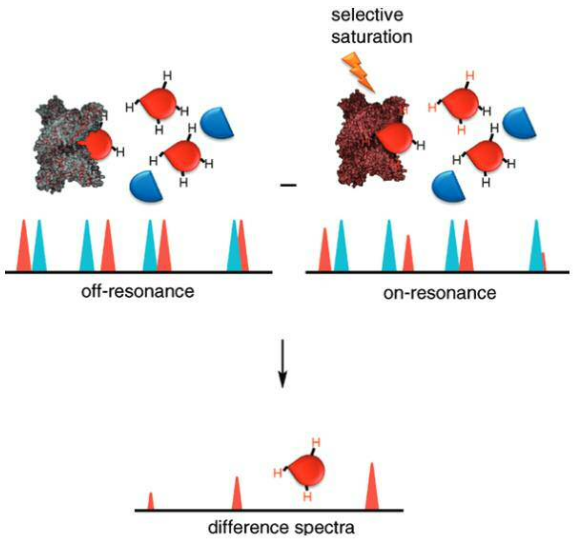


Fig. 3.29 The basis of Saturation Transfer Difference-NMR.

In this Thesis, STD experiments were performed for an OS/protein molar ratio of approximately 30:1. Protein concentration was determined using UV absorbance at 280 nm and an extinction coefficient $E^{0.1\%}$ of 3 and 1.4, for VAA and RCA. Selective saturation of the protein was achieved by using a train of 40 Gaussian-shaped pulses of 50-ms each, separated by a 1 ms delay (approx. saturation time of 2 s). The residual HDO signal was suppressed by hard-pulse gradient tailored excitation (WATERGATE) and background protein resonance signals were subtracted. An *off-resonance* frequency of $\delta = 100$ ppm and *on-resonance* frequency of either $\delta = -0.5$ or 7 ppm were applied, targeting a spectrum region where no signals are observed for the ligand. For the epitope mapping analysis, a quantitative analysis of the STD signals was carried out, the resulting STD percentages being given in Tables 4.5 and 4.6 in section 4.1.6.3.2.

3.9 Bacteria binding and functional assays in solution

Assays with live NTHi bacteria were performed at the laboratory of Dr. Junkal Garmendia in biosafety level 2 installations.

3.9.1 Galectin binding to bacteria at exponential growth phase

Bacteria were grown on chocolate-agar plates as described in section 3.1. One hundred and 150 µl of NTHi culture in exponential growth phase was centrifuged for 10 min at 18,000×g and 4 ºC. Pellets were washed with PBS, resuspended in 50 µl PBS containing the galectin of interest in the absence or presence of 50 mM lactose, and incubated for 45 min at 37 ºC and 5% CO_2. Control inoculate suspensions in the absence of galectins were also prepared. Suspensions were then centrifuged for 5 min at 6,000×g, and the supernatants were mixed with SDS-PAGE loading buffer (section 3.11.1). Pellets were washed once in PBS, centrifuged, supernatant was withdrawn and the pellet resuspended directly in loading buffer. All samples (supernatants and pellets) were heated for 5 min at 100 ºC, and loaded in gels containing 12% total acrylamide. Gels were stained with Coomasie Blue and scanned with a GS-800™ Calibrated Densitometer (BioRad). Images were analysed with Quantity One® 1-D analysis software (BioRad) for identifying and quantifying the bands corresponding to galectins.

3.9.2 Bacterial susceptibility assay

NTHi susceptibility to galectins was determined as previously described (Llobet *et al.*, 2011; Morey *et al.*, 2013) with slight modifications. Bacteria were inoculated in sBHI

medium, grown as described in section 3.1, and harvested (2,600×g for 20 min at 20 ºC) when the exponential growth phase was reached (OD_{600} = 0.3). The pellet was washed with PBS and subsequently resuspended in PBS at OD_{600} = 0.15 (corresponding to approximately 10^8 cfu/ml, and serial dilutions in PBS were prepared up to 10^{-3} dilution. To 25 µl of galectin solution in either PBS or PBS_{β}, 5 µl of bacterial 10^{-3} dilution was added and mixed by soft pipetting. The mixture was incubated at 37 ºC and the whole 30 µl was plated on sBHI-agar at appropriate time intervals (0, 15, 30, 45, 60 and 90 min). PBS_{β} was included in control preparations (absence of galectins) to evaluate potential effects of *β*-mercaptoethanol on the bacterial viability. Colony counts were determined and the results were expressed as percentage of the colony count of bacteria not exposed to galectins. All measurements were performed with duplicate samples in at least three independent experiments.

3.9.3 Epithelial adhesion/invasion assay

A549 human alveolar basal epithelial carcinoma cells (immortalised type II pneumocytes, ATCC CCL-185) were maintained in Roswell Park Memorial Institute tissue culture medium (RPMI-1640, Sigma) supplemented with 10 mM HEPES (Biowest), 10% heat-inactivated foetal bovine serum (FBS, Sigma) and a combination of antibacterial/antifungal drugs (100 IU/ml penicillin, 0.25 µg/ml amphotericin B, and 0.1 mg/ml streptomycin sulphate, Sigma) in 25 cm^2 tissue culture flasks at 37 ºC in a humidified 5% CO_2-containing atmosphere, as previously described (Morey *et al.,* 2011). Cells were seeded in 24-well plates to reach 4×10^5 cells/well and serum-starved for 16 h before infection by replacement of medium with supplemented RPMI-1640 lacking FBS. A confluence of 90% was reached at the time of bacterial infection.

For NTHi infection (Fig. 3.30), bacteria grown for 16 h (stationary phase) were recovered from a chocolate-agar plate with 1 ml PBS, and the bacterial suspension was adjusted to OD_{600} = 1 (*input*, corresponding to ~10^9 cfu/ml). For the cell-adhesion assay, A459 cells were infected in 500 µl Earle's Balanced Salt Solution (EBSS, Life Technologies) for 30 min at 37 ºC and 5% CO_2 with 50 µl of *input* to get a multiplicity of infection (MOI) of 100:1 (100 bacteria per cell). Experiments were performed in triplicate, in the absence or presence of human galectins at 5 µM, after exhaustive dialysis against PBS/PBS_{β}. Three experimental modes were followed: direct infection in galectin-containing EBSS, infection after 60 min preincubation of bacteria with galectins, and infection after 60 min preincubation of cells with galectins. After 30 min,

cells were washed thrice with PBS and subsequently lysed with 300 µl of PBS containing 0.025% w/v saponin (Sigma), for 10 min at 20 ºC. Upon cell lysis, the *output* was recovered and serially diluted. One hundred µl of 10^{-4} and 10^{-3} dilutions were plated on sBHI-agar and the number of cfu counted. Bacterial adhesion to the epithelial

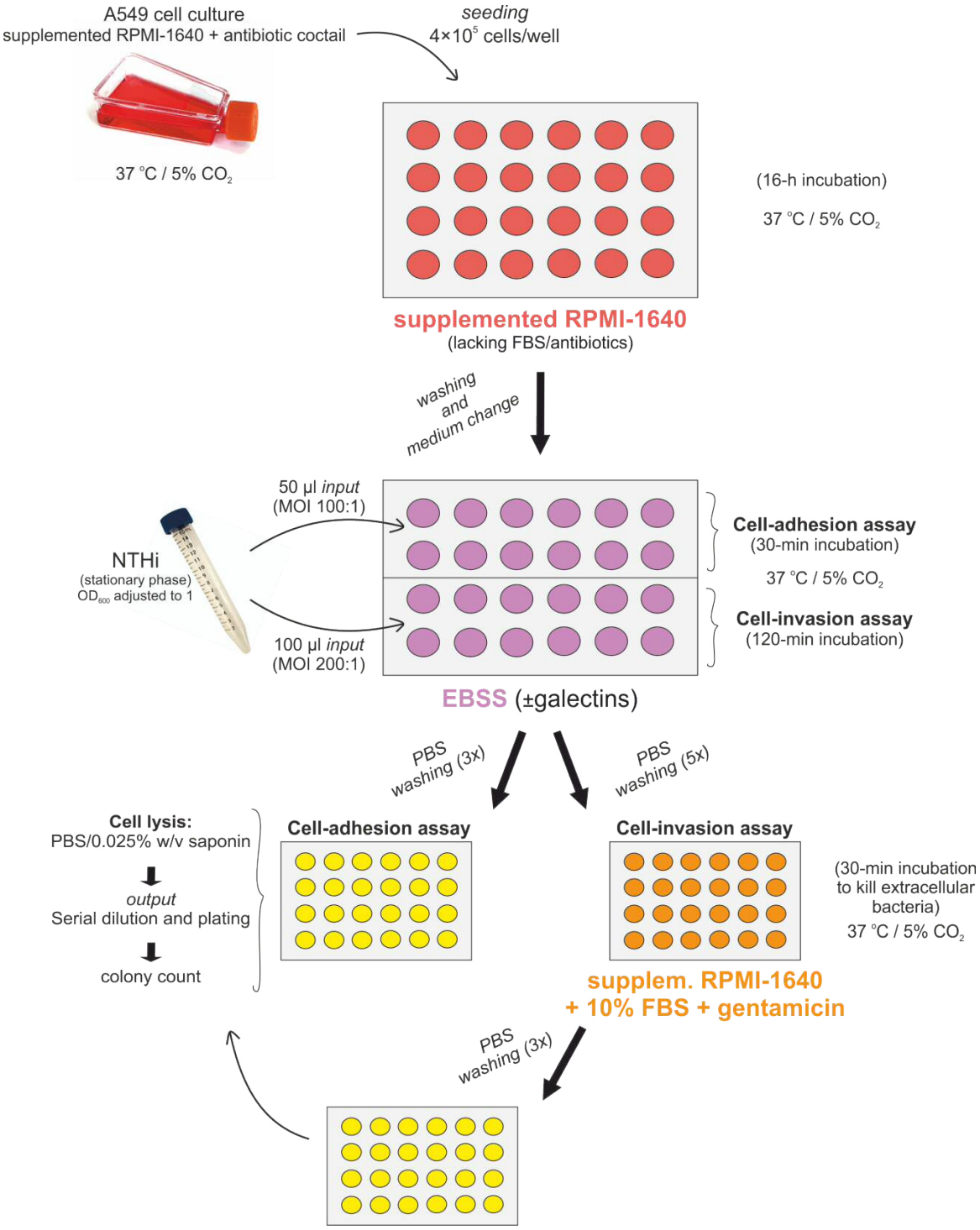


Fig. 3.30 Epithelial NTHi adhesion and invasion assays in presence/absence of galectins. For the adhesion assays, three experimental modes were followed as described in the text. Upon serial dilution and plating of the output, adhesion and invasion were represented as percentage of c.f.u. per well, taking the input as 100%.

cells was represented as the percentage of cfu per well, taking the *input* as 100%.

For the bacterial cell-invasion assay (Fig. 3.30), A459 cells were infected for 2 h as described above with 100 µl of *input* (MOI 200:1), in the absence or presence of galectins. This is the time considered necessary for bacterial internalisation upon adhesion. Then, cells were washed 5 times with PBS and incubated for another 30 min with RPMI-1640 containing 10 mM HEPES, 10% FBS, and 200 µg/ml gentamicin (Sigma) to kill extracellular bacteria. Cells were then washed thrice with PBS and lysed as described above. The *output* was serially diluted and 100 µl of 10^{-2} and 10^{-1} dilutions were plated on sBHI-agar. Results were represented again as the percentage of recovered cfu per well.

3.9.4 Galectin binding visualisation by confocal microscopy

NTHi375 bacteria were harvested at stationary phase and fixed with paraformaldehyde, washed and their nucleic acids were stained with SYTO® 13, all as described in section 3.1. For lectin-binding, bacteria suspensions (OD_{600}=0.5) were incubated for 60 min at 37 °C, in presence or absence of biotinylated hGal-1 or his-tagged galectins (hGal-3, hGal-4 and hGal-8) at 10 µM. Meanwhile, a mouse anti-his Ab and a secondary, biotinylated anti-mouse Ab were precomplexed in solution at 1 and 3 µg/ml respectively. Following galectin incubation, preparations were centrifuged and washed thrice with PBS, then incubated for 45 min at 37 °C with the antibody complex (for hGal-3, -4 and -8) or PBS (for hGal-1-biotin). After washing, they were incubated with 1 µg/ml streptavidin-AlexaFluor® 647 for 30 min, washed and tested in the microscope.

Images were collected with a Leica TCS SP5 Acousto Optical Beam Splitter confocal microscope (Mannheim, Germany) with 20X and 63X oil immersion optics. Laser lines at 488nm and 561nm for excitation of SYTO® 13 and AlexaFluor® 647 were provided by an Ar laser (100 Mw) and a DiodeP Solid State laser (10 Mw). Detection ranges were set to eliminate crosstalk between fluorophores.

3.10 Affinity chromatography

The isolation and purification of lectins by affinity chromatography exploits the unique property of these proteins of binding saccharides specifically and reversibly. The lectin to be purified is passed through a column containing an insoluble matrix to which a specific ligand has been covalently attached (Cuatrecasas, 1970; Vretblad, 1976).

Proteins not exhibiting appreciable affinity for the ligand elute through the column, whereas those bound to the functionalised matrix can be later dissociated by competitive elution with an appropriate hapten sugar. In this Thesis, affinity chromatography has been used to compare functionality among different batches of ConA and recombinant human galectins, as well as for the purification of their active fractions.

Affinity chromatography was carried out in 5-ml columns containing either commercial mannan-agarose (Sigma) or fetuin-Sepharose 4B prepared as follows. Cyanogen bromide (CNBr)-activated Sepharose 4B (GE Healthcare) was initially suspended at 0.5% (w/v) in 0.1 M HCl and incubated for 1 h at 20 ºC with sporadic shaking, to eliminate possible imidocarbonate groups that could interfere during the chromatographic procedure (Tercero and Díaz-Mauriño, 1988). The gel was next separated by filtration and washed first with abundant milli-Q water and afterwards with a small volume of 0.1 M $NaHCO_3$, pH 8.3, 1 M NaCl. The gel was immediately mixed with an equal volume of a fetuin solution (5-10 mg/ml) in the same buffer, and the suspension was incubated overnight at 4 ºC, under soft spinning. After incubation, the gel was separated from the supernatant by filtration and washed with 5 volumes of the bicarbonate buffer. The fetuin-Sepharose matrix was resuspended in 1 M ethanolamine, pH 9, in order to de-activate any remaining active cyanide groups, and incubated for 2 h at 20 ºC, under soft spinning. Finally, the gel was washed with PBS. The amount of immobilised fetuin (2.5-4.5 mg/ml) was calculated from the difference between the initial quantity of protein added and the amount of fetuin determined in the supernatant and washes by measuring the absorbance at 280 nm ($E_{1cm}^{0.1\%} = 0.494$).

The mannan-agarose and fetuin-Sepharose 4B columns were equilibrated at room temperature with PBS or PBS$_{\beta}$, and a 100 µl sample of the lectin (1 mg/ml) in the same buffer was injected. Elution was monitored by measuring absorbance at 280 nm with a UV Uvicord S2 (LKB) detector and the signal was recorded using a Shimadzu C-R3A integrator. The lectin retained in the column was finally eluted with running buffer containing 0.1-0.2 M competing sugar. Flow rate was fixed at 0.25 ml/min using a peristaltic LKB Pump P1 (Pharmacia).

3.11 Colorimetric methods

3.11.1 Determination of protein concentration

Protein concentration was determined colorimetrically in 96-well polystyrene microtiter plates (Costar) following the Bradford method (Bradford, 1976). This method is based on the binding of the dye Coomasie Blue G-250 (red colour in the free-state, maximum absorption at 465 nm) to the proteins, with the formation of a blue-coloured protein-dye complex of high extinction coefficient (maximum absorption at 595 nm). The binding occurs rapidly (it takes approximately 2 min) and the blue complex remains dispersed in solution and stable for a relatively long time (approximately 1 h). The assay is highly reproducible in most cases, no interference of cations or sugars is expected, and it can be run in the presence of most buffers by including appropriate controls.

Protein solutions containing 2 to 8 µg of protein in a 120 µl volume were pipetted in triplicate in the microplate wells, and 80 µl of diluted Bradford Reagent (Bio-Rad®, 1:1 dilution with milli-Q water) was added. The Bradford Reagent contains 100 mg of Coomasie Blue G-250 dissolved in a mix of 100 ml phosphoric acid at 85% and 50 ml of ethanol at 95%. The plate was shaked and incubated for 10 min at 20 ºC and the absorbance was measured at 595 nm, using a Bio-Rad® 3550 plate-reader. Standard curves built with triplicate samples of bovine serum albumin at different concentrations permitted interpolation of the absorbance values and calculation of the concentration of the protein sample.

3.11.2 Determination of hexose concentration

Hexose concentration was determined colorimetrically in 96-well polystyrene microtiter plates (Costar) following the phenol-sulphuric acid method (Dubois *et al.*, 1951). The assay is based on the formation of a coloured, stable phenolic complex (yellow-orange), as a result of the reaction between phenol and the carbohydrates after dehydration of hydrolysed saccharides to furfural derivatives due to the energetic addition of concentrated sulphuric acid, what generates heat.

Sugar solutions containing 0.5 to 10 µg of sugar in a 25 µl volume were pipetted in triplicate into the microplate wells, and 25 µl of 4% phenol in milli-Q water was added and mixed thoroughly, followed by the energetic addition of 200 µl of concentrated H_2SO_4. The plate was shaken and incubated for 1 h at 20 ºC and the absorbance was

measured at 490 nm, using a Bio-Rad® 3550 plate-reader. Standard curves built with triplicate samples of the appropriate sugar at different concentrations permitted interpolation of the absorbance values and calculation of the sugar concentration per well/sample.

3.11.3 Quantitation of bacterial lipopolysaccharides by the Purpald® assay

The molar concentration of LPS and LOS solutions was determined colorimetrically by the Purpald® assay (Lee and Tsai, 1999) in 96-well polystyrene microtiter plates (Costar). The method employs the oxidation by periodate, at room temperature, of unsubstituted terminal vicinal glycol (UTVG) groups present in 2-keto-3-deoxy-D-manno-octonate (Kdo) and L-glycero-D-manno-heptose (Hep) yielding formaldehyde.

Formaldehyde reacts with Purpald® giving a colourless adduct, which is further oxidised by periodate to form the end product of purple colour, as shown in Fig. 3.31. This assay allows the quantitation of many types of LPS, since they just need to possess at least one UTVG. This minimal requisite makes possible the estimation of the molar concentration of the NTHi LOS, which cannot be detected by the typical Kdo assay (Osborn, 1963) due to substitutions at the C-4 and C-5 positions of Kdo.

Duplicates of LPS/LOS solutions in milli-Q water at different concentrations, prepared by serial dilution, were pipetted in 96-well polystyrene microtiter plates at a volume of 50 µl. In each well, 50 µl of 32 mM $NaIO_4$ was added, and the plate was incubated for

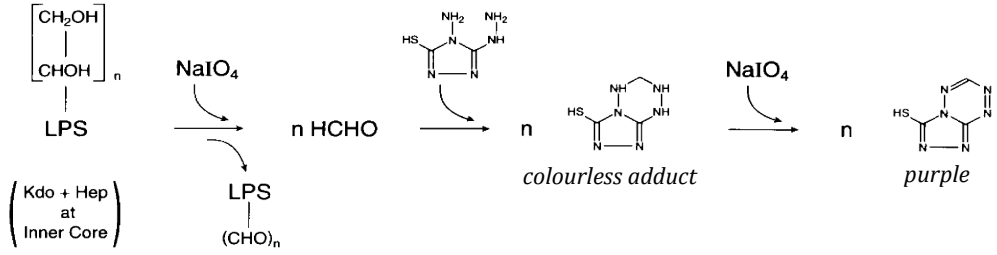


Fig. 3.31 Chemistry of the purpald assay for LPS/LOS quantitation. The unsubstituted terminal vicinal glycol of the Kdo and heptose residues in LPS molecules is oxidised by periodate, producing formaldehyde (n molecules corresponding to those of glycol) that reacts with purpald and yields a colorless adduct, further oxidised by periodate to form the end product of purple color. Adapted from (Lee and Tsai, 1999)

25 min, followed by addition of 50 µl of 136 mM Purpald® (Sigma-Aldrich) in 2 N NaOH. After further incubation for 20 min, 50 µl of 64 mM $NaIO_4$ was added per well and the plate was incubated for another 20 min. All reactions took place at room

temperature. The foam generated during the last step was eliminated by addition of 20 µl of 2-propanol. Absorbance was measured at 550 nm, using a SpectraMax® Plus 384 plate-reader located at Instituto de Química Orgánica General, Centro de Química Orgánica "Lora Tamayo" (CSIC, Madrid, Spain). Standard curves of duplicate samples of Kdo, Hep, formaldehyde and 5-N-acetyl-neuraminic acid (Neu5Ac) permitted the interpolation of the absorbance values and calculation of the LPS/LOS molar concentration, based on the theoretical UTVG number per molecule.

3.12 Polyacrylamide gel electrophoresis (PAGE)

Polyacrylamide gel electrophoresis (PAGE) is an analytical technique widely used for the separation of biological macromolecules with different electrophoretic mobility. The mobility is a function of the size, conformation and charge of the biomolecule. These parameters can be controlled by addition of anionic detergents (sodium dodecyl sulphate (SDS), sodium deoxycholate (DOC), etc.), reducing agents (dithiothreitol or *β*-mercaptoethanol) and other factors (heating, addition of urea, etc.).

3.12.1 Protein polyacrylamide gel electrophoresis (SDS-PAGE)

Protein polyacrylamide gel electrophoresis in the presence of SDS (SDS-PAGE) was carried out following the Laemmli protocol (Laemmli, 1970), using a Mighty Small II SE 250 system (Hoefer Scientific Instruments) associated with a source of variable voltage and amperage (EPS 500/400, Pharmacia). Proteins covered with SDS have a relatively equal negative charge and their electrophoretic mobility depends mainly on their size.

In order to obtain efficient resolution in the range between 12 and 55 kDa, polyacrylamide separating gels were prepared at a 15% total acrylamide concentration (%T, total concentration of both acrylamide and bisacrylamide). The concentration of bisacrylamide (%C, cross-linker) was 3% of the total acrylamide concentration. Stacking gels contained 4% T and 3% C. Protein samples were mixed 1:1 with sample buffer, containing 0.05 M Tris-HCl, pH 6.8, 10% glycerol (v/v), 10% SDS (w/v) and 0.1% (w/v) bromophenol blue, and incubated for 3 minutes at 100 °C before loaded into the gel. In the case of electrophoresis under reducing conditions, the buffer also contained 0.2 M *β*-mercaptoethanol. The running buffer was Tris-glycine-SDS buffer, containing 25 mM Tris, 192 mM glycine and 0.1% SDS (w/v), and the electrophoresis was carried out under constant amperage (25 mA). Afterwards, proteins were stained

with Coomassie Brilliant Blue R-250 (Fluka) at 0.25% (w/v) in methanol/acetic acid/water (45:5:50 v/v/v). A protein-standard ladder was used for monitoring the electrophoretic mobility of the samples in the low molecular weight range (Bio-Rad). It contained phosphorylase b (97.4 kDa), BSA (66.2 kDa), ovalbumin (45 kDa), carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa) and lysozyme (14.4 kDa).

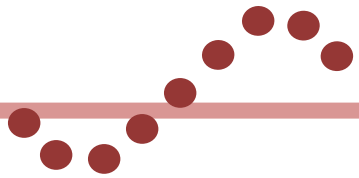
In the case of electrophoresis of bacterial pellets (see section 3.9.1) separating gels were prepared at 12% total acrylamide concentration, of which 3% was bisacrylamide. Stacking gels contained again 4% T and 3% C, while the loading and running buffers were prepared as described above. The prestained 12-protein standard ladder used was Protein Marker VI, ranging 10-245 kDa (Panreac). These electrophoreses were run at the laboratory of Dr. Junkal Garmendia.

3.12.2 Lipopolysaccharide polyacrylamide gel electrophoresis (DOC-PAGE)

DOC-PAGE has been selected as the method of choice for the analysis and quantitation of NTHi LOS, which is comparable in size to the LPS rough form of other Gram-negative bacteria. A novel discontinuous Tricine-DOC-PAGE protocol was developed and optimised. The system consisted in a separating gel containing 16.5% T, 6% C and 6 M urea, overlaid by a spacer gel containing 10% T and 3% C, and an upper stacking gel with an even bigger pore size (4% T, 3% C) for alignment of all LOS species before entering the spacer gel. All gels contained a final DOC concentration of 0.15%, below the critical micelle concentration of 0.16%. Five-µl samples were mixed 1:1 with sample buffer containing 1% DOC (w/v), 0.1 M Tris-HCl, pH 6.8, 20% glycerol and 0.1% bromophenol blue (w/v), and directly loaded into the gels without any previous incubation. Electrophoresis was performed with two distinct buffers for anode and cathode; the anode buffer contained 0.2 M Tris-HCl, pH 8.9, whereas the cathode buffer contained 0.1 M Tris Base, 0.1 M Tricine and 0.1% SDS (w/v), pH 8.25. As molecular weight marker, the S-form of *Salmonella enterica* s. minnesota (Sigma) was selected.

LPS/LOS bands were detected by silver staining. To this aim, a sensitive protocol combining sugar oxidation and gel sensitisation was optimised. In detail, polyacrylamide gels were fixed for 45 min in 100 ml of fixation solution containing methanol/acetic acid/milli-Q water (40:10:50 v/v/v). Upon completion of the fixation procedure, further 50 ml of fixation solution containing 1.05 g of H_5IO_6 was added, and incubation was continued for another 10 min. Then, the gel was thoroughly washed with

milli-Q water, 6 times for 5 min each, and the sensitising agent $Na_2S_2O_3$ was added at a concentration of 0.2 g/l. After 1 min with stirring, the gel was washed with milli-Q water twice for 1 min, to remove the excess of ions, and then incubated at 4^oC for 45 min with staining reagent, containing 0.1% $AgNO_3$ (w/v) and 0.028% formaldehyde. All the described incubation and washing steps were carried out under continuous shaking. Following a 30-s wash with milli-Q water, alkaline developer containing 3% K_2CO_3 (w/v), 0.05% formalin and 10 mg/l $Na_2S_2O_3$ was added. Once LPS bands were satisfactorily revealed (usually after 5-10 min) the reaction was stopped by addition of 5% acetic acid. Finally, gels were thoroughly washed with milli-Q water for at least 30 min and LPS bands were quantitated by densitometry using the UNSCAN-it-gel v6.1 software (Silk scientific), after scanning the gels with a Perfection 3200 Photo scanner (EPSON).



RESULTS

We cannot solve our problems with the same
thinking we used when we created them.

Albert Einstein

4.1 Glycosylation patterns of pathogenic bacteria as self and non-self markers for endogenous lectins

Bacterial surfaces are coated with a variety of carbohydrate-rich structures that usually serve to typify strains (Peterson and Quie, 1981; Hsu *et al.*, 2006; Kilcoyne *et al.*, 2014). The accessibility and density of glycans on the bacterial surface make them key targets for recognition by host lectins and/or soluble molecules, for triggering defence responses or as mechanism exploited by the pathogen for both immune escape and adherence to, and invasion of host cells (Sahly *et al.*, 2008; Davicino *et al.*, 2011; Kumar *et al.*, 2011). In this Thesis, a combined designer' microarray and QCM approach was developed for the analysis of carbohydrate-mediated interactions directly on the bacterial surface, thus preserving the native environment, distribution and vicinity of the glycan targets (Kalograiaki *et al.*, 2016).

Bacteria microarrays were previously developed in our group using *Klebsiella pneumoniae* as model (Campanero-Rhodes *et al.*, 2015). This is a carbohydrate-rich Gram-negative bacterium that displays a polysaccharidic capsule and an O-chain-containing LPS on its surface. Here we have examined the usefulness of the bacteria microarray approach for the study of other bacteria displaying a different repertoire of surface carbohydrate structures. This is the case of nontypeable *Haemophilus influenzae* (NTHi), which is a common source of upper respiratory infection, primary causative agent of pediatric otitis media, and also associated with chronic obstructive pulmonary disease (COPD) exacerbation in adult patients (Agrawal and Murphy, 2011). NTHi is not capsulated and displays a much simpler LPS that does not contain repetitive O-antigen, called LOS (described in section 1.2). LOSs are major virulence factors and are characterised by the capacity of incorporating host-"mimicking" epitopes, thereby offering pathogens an advantage in their survival and interplay with the host. Several other important pathogens of the upper respiratory tract, including *Neisseria*, *Moraxella* and *Bordetella* spp., also present LOSs with oligosaccharide structures limited to 10 units (Preston *et al.*, 1996). Thus, NTHi serves as archetype for exploring the applicability of bacteria microarrays to a relevant class of human pathogens. In addition, novel QCM bacteria-based biosensors were developed, also using NTHi as model.

The glycosylation patterns of different NTHi strains were examined in the microarray set-up using a panel of lectins with diverse carbohydrate-binding specificities. In

addition, binding kinetics for representative NTHi-lectin pairs was investigated using the newly developed QCM bacteria-based chips. The LOS was shown to be a target for some lectins, and in two selected cases the recognized epitopes were defined at atomic level using NMR-STD experiments. Moreover, the recognition of NTHi by different innate immunity lectins was explored. Finally, a pilot study of the influence of selected endogenous lectins on the microbial viability *in vitro* and on the bacterial potential for adhesion and invasion in cell culture, was carried out.

4.1.1 Preparation and validation of NTHi arrays

For the initial preparation and validation of NTHi microarrays, the NTHi strain 375 (hereafter referred to as NTHi375), a clinical isolate from otitis media, was selected. This strain is highly transformable (Maughan and Redfield, 2009), its whole genome and methylome is available (Mell *et al.*, 2014), and its LOS structure (Fig. 4.1) has been determined by mass spectroscopy (Bouchet *et al.*, 2003). More important, a library of isogenic mutant strains bearing disruption of single genes involved in LOS biosynthesis, therefore expressing sequentially truncated LOS, was available (Morey *et al.*, 2013), what should facilitate the identification of epitopes recognised by the tested lectins.

As illustrated in Fig. 4.1, genes disrupted in the mutant library included *opsX*, encoding a heptosyltransferase adding the first L-glycero-D-manno-heptose to Kdo (Gronow *et*

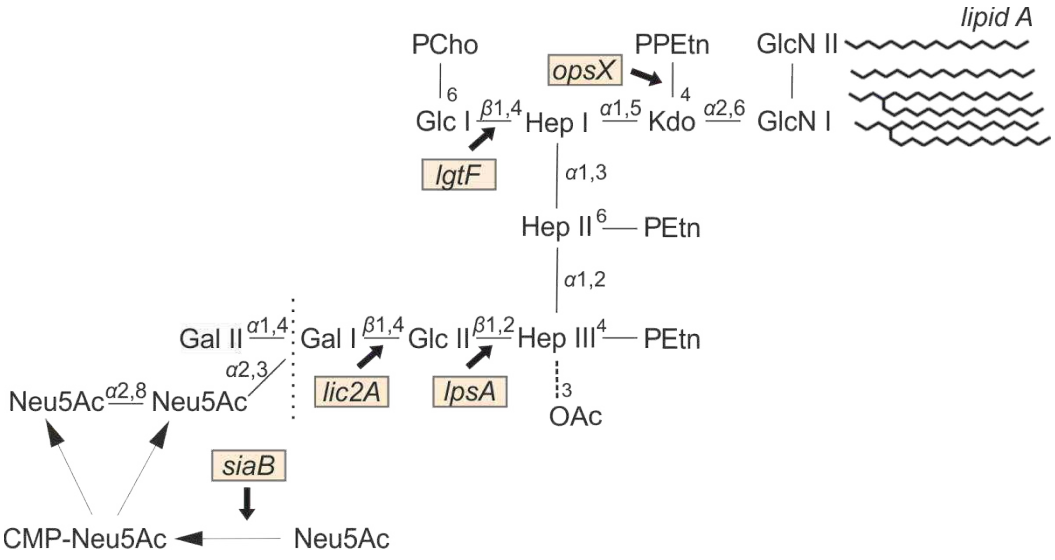


Fig. 4.1 LOS structure proposed for NTHi375 (Bouchet *et al.*, 2003)
Genes of interest are highlighted in light pink, with arrows indicating their point of action. The major glycoform (Hex4) bears the Hep inner core and four Hex moieties with Gal II at the non-reducing end. Sialylated species, with one or two Neu5Ac residues, represent a total of ~14% (Fox *et al.*, 2006). Abbreviations: GlcN, glucosamine; PEtn, phosphoethanolamine; Hep, heptose; Glc, glucose; Gal, galactose; PCho, phosphorylcholine; OAc, O-acetyl; P, phosphate.

al., 2005); *lgtF* and *lpsA*, encoding glycosyltransferases responsible for adding Glc I and II to Hep I and III respectively (Hood *et al.*, 2004); and *lic2A*, a phase-variation regulated gene encoding a glycosyltransferase that adds Gal as part of a lactose epitope. Of note, NTHi can capture host-derived 5-acetyl-neuraminic acid (Neu5Ac) from its media and incorporate it into its LOS as terminal sugar residue (Bouchet *et al.*, 2003), mono- and di-sialylated species having been reported to represent 9% and 5% respectively of NTHi375 LOS species (Fox *et al.*, 2006). The *siaB* gene encodes a synthetase that catalyses the formation of CMP-Neu5Ac, the nucleotide sugar donor used by NTHi sialyltransferases (Hood *et al.*, 1999). Consequently, the *siaB* gene-disrupted mutant does not display sialylated carbohydrate structures. For mutant designation, the symbol Δ is used to indicate disruption of the respective gene, *e.g.* Δ*siaB*.

NTHi375 and selected mutant strains were fluorescently labelled with SYTO-13® and printed onto nitrocellulose-coated glass slides as triplicates at four different dilutions. Fluorescence signals of printed bacteria were reliably detectable, and showed the expected concentration-dependent trend (Fig. 4.2, left panel).

For validation of the NTHi microarrays, the binding of a rabbit polyclonal anti-NTHi serum, raised against a mixture of acetone-killed NTHi strains 375, 398 and 2019

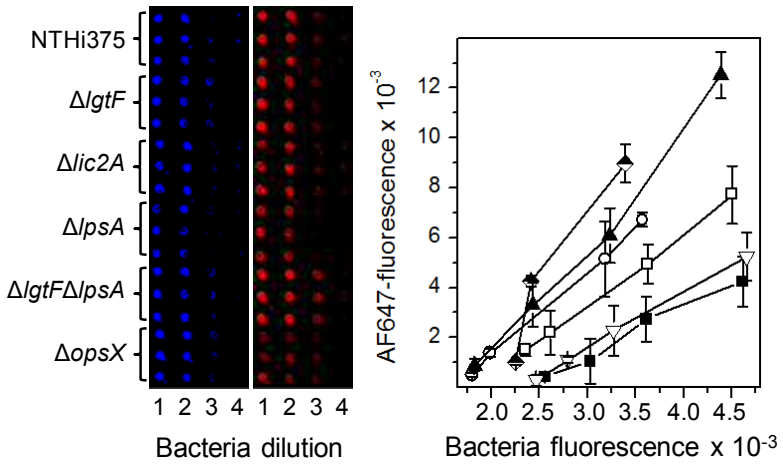


Fig. 4.2 Binding of rabbit polyclonal anti-NTHi serum to NTHi arrays
Left panel, Slides scanned for SYTO-13® fluorescence: bacteria were printed in triplicate at concentrations 1 to 4, corresponding to OD_{600} of 1, 0.6, 0.3 and 0.1, respectively. Middle panel, Slides scanned after overlay for AF647 fluorescence. Right panel, normalised binding intensity: wild type NTHi375 (□), and mutants Δ*lgtF* (▲), Δ*lic2A* (○), Δ*lpsA* (▽), Δ*lgtF*Δ*lpsA* (◆), and Δ*opsX* (■). Fluorescence signals of printed bacteria were normalized to wild type NTHi375, taking into account the labelling efficiency. Data shown correspond to the mean of two different experiments and error bars indicate the standard deviation. Fluorescence intensity values are expressed in relative fluorescence units (rfu).

(Morey *et al.*, 2011), was examined. All strains tested were bound by the anti-NTHi serum in a dose-dependent manner, proving accessibility of the recognised epitopes (Figure 4.2, middle and right panels).

Worth to mention, truncation of the Hep I ramification of the LOS (Δ*lgtF*) alone, or with simultaneous truncation of the Hep III ramification (Δ*lpsA*), seemed to favour recognition, suggesting an increase in epitope(s) exposure. Although the precise epitopes recognized by the anti-NTHi serum are not known, it seems likely that recognition of exposed domains of the immunogenic outer membrane proteins P1, P5, HMW or Hap could be enhanced in these mutant strains (Winter and Barenkamp, 2003; Euba *et al.*, 2015; Tchoupa *et al.*, 2015). Thus, having proved the efficiency of NTHi microarrays for detecting binding of soluble molecules, the way was paved to the analysis of bacterial surface features, as e.g. availability of carbohydrate epitopes for lectin recognition.

4.1.2 Profiling of accessible glycan chains on the surface of NTHi375

To explore the sugar presence and accessibility on the surface of the printed bacteria, we first examined the binding to NTHi375 of a panel of 19 biotinylated lectins of known binding specificities (see Table 3.2 for details). Notable differences in the binding were found (Fig. 4.3, upper panel), evidencing the unavailability of ligands with the appropriate structure/accessibility for meaningful recognition by some of the tested lectins, such as e.g. LEL, PSA or HHL. This was supported by the fact that these lectins strongly bound to the proper control glycoproteins also printed in the arrays but not to non-glycosylated RNase A, thus proving out to be functional.

For all of the lectins, the specificity of the binding was assessed in parallel assays carried out in the presence of the suitable hapten for each lectin (Table 3.2). Again, noteworthy differences in the inhibitory potential of the tested sugars were observed (Fig. 4.3, upper panel). No inhibition was observed for DSL or LPA, and only low inhibition was detected for GSL-IB4, GSL-I, PNA, HPA or GNA. This behaviour could derive from a very strong carbohydrate-mediated binding of these lectins to components of the bacterial surface, including the peptidoglycan (terminal GlcNAc) (Burroughs *et al.*, 1993), but it could also reflect the occurrence of non-carbohydrate-mediated interactions. In contrast, the binding of 9 lectins, namely VAA, RCA, MAL-I, SBA, WFA, WGA, PSqL, SNA and ConA, was inhibited to a moderate/large extent. As

consequence, the pattern of hapten-inhibitable lectin binding (Fig. 4.3, lower panel) clearly differed from that of total binding. The results unambiguously revealed the presence of carbohydrate structures on the bacterial surface specifically recognised by the lectins, hinting at the presence of certain epitopes as described below.

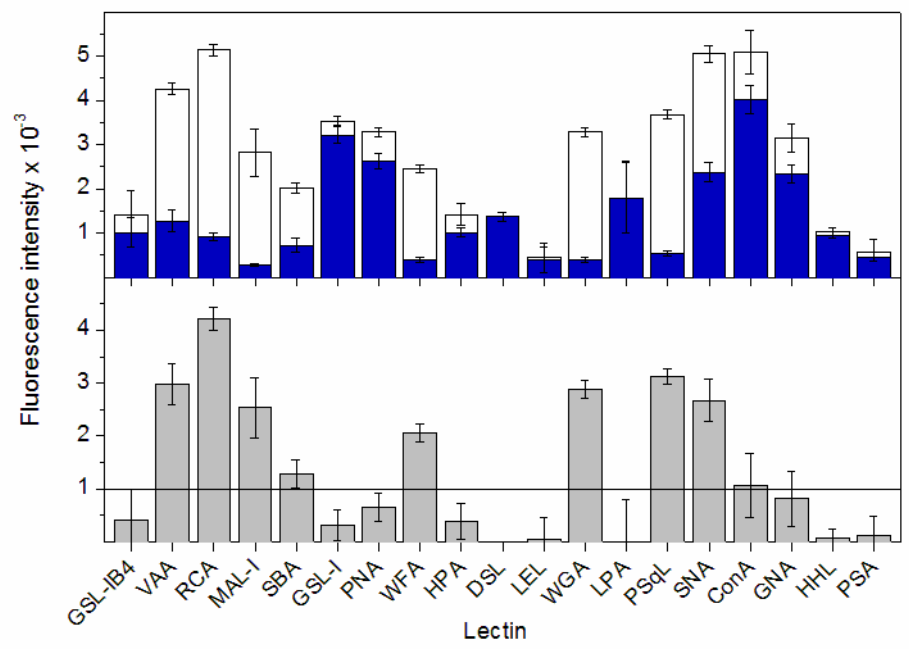


Fig. 4.3 Binding of reference lectins to NTHi375
NTHi375 was printed in triplicates at concentrations corresponding to OD_{600} = 1 and binding of biotinylated lectins was assessed by incubation with AF647-streptavidin. Upper panel: Lectin binding to the NTHi arrays in the absence (white) and presence (blue) of the respective lectin hapten (Table 3.2). Lower panel: hapten-inhibitable binding calculated from data shown in the upper panel. The horizontal line marks the threshold for significant specific binding (fluorescence intensity 1×10^3 rfu). Fluorescence intensity values are expressed in rfu. Data shown correspond to the mean of at least two different experiments. Error bars indicate the standard deviation of the mean.

The noticeable binding of the Gal-specific lectins VAA and RCA evidenced the availability of Gal-containing structures, and binding of MAL-I pointed to the presence of the lactose epitope (Geisler and Jarvis, 2011), which could also account, at least in part, for WFA binding (Kurokawa *et al.*, 1976). Finally, the sialic acid-binding WGA, PSqL and SNA revealed the availability of Neu5Ac for lectin binding (Monsigny *et al.*, 1980; Shibuya *et al.*, 1987; Mo *et al.*, 2000). Of note, both Gal and Neu5Ac residues are integral constituents of NTHi375 LOS (Fig. 4.1), which also contains Glc and L,D-hepto-mannose residues, suggested to serve as docking points for ConA (Goldstein *et al.*, 1965; Baenziger, 1985; Jaipuri *et al.*, 2008). Hence, the LOS could be a prevalent target for these lectins, prompting us to analyse their binding to our library of mutant strains expressing a sequentially truncated LOS.

4.1.3 Glycotyping of NTHi375 mutants: the effect of LOS truncation

Those lectins that exhibited hapten-inhibitable binding to NTHi375, in terms of fluorescence intensity, of 1×10^3 rfu or above (denoted by the horizontal line in Fig. 4.3, lower panel) were selected for extended study. As illustrated in Fig. 4.4, LOS truncation had differential consequences on lectin binding. The shortest truncations, as those found in Δ*siaB*, Δ*lgtF* and Δ*lic2A* strains, had disparate lectin-specific consequences, with both reduced and enhanced binding compared to NTHi375 (Fig. 4.4 and Table 4.1), whereas deletion of the Hep III ramification (Δ*lpsA*) resulted in decreased binding for many lectins, and the most truncated mutants, lacking the outer core (Δ*lgtF*Δ*lpsA*) or even all core sugars (Δ*opsX*) showed a moderate to strong drop in the binding of most lectins.

A close view to each mutant background and the lectins' specificities helped to explain the binding patterns here revealed. Starting with NTHi375Δ*siaB*, this mutant strain does not express the enzyme responsible for the synthesis of the CMP-Neu5Ac donor. Consequently, its LOS, as well as any other carbohydrate structure of the bacteria, cannot be sialylated. Consistently, the binding of the Neu5Ac-specific lectins WGA, PSqL and SNA decreased significantly. The capability of these lectins to recognise other sugar epitopes, as e.g. GlcNAc for WGA (Monsigny *et al.*, 1980) or lactose for PSqL (Mo *et al.*, 2000) and SNA (Nsimba-Lubaki *et al.*, 1986), could account for their remaining binding to the sialylation-defective mutant. Contrariwise, the absence of Neu5Ac residues masking Gal or GalNAc groups might explain the significant increase

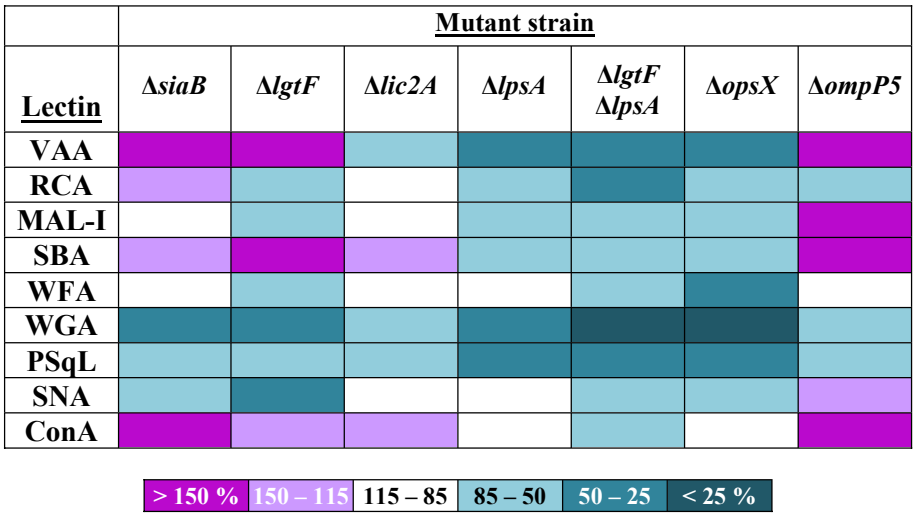


Fig. 4.4 Relative binding strength of selected reference lectins to NTHi375 wild type and mutant strains. Binding of biotinylated lectins to NTHi375 wild type and mutant strains printed onto NC-coated glass slides was assessed simultaneously, and the relative binding strength was calculated as percentage taking for each lectin the binding to wild type NTHi375 as 100%. Detailed data are given in Table 4.1.

in the binding of SBA, RCA and VAA. A similar shielding effect of sialic acid on the binding of SBA was repeatedly reported in cell binding and histochemical studies (Takahashi, 1992; Abe *et al.*, 1996). Moreover, it is considered that substitutions at position 3´ of Gal, as those found in the minor, α(2,3)-sialylated isoform of NTHi375 LOS, may abolish RCA and VAA binding because this is a key position for Gal recognition (Solís *et al.*, 1993; Jiménez *et al.*, 2008). In contrast, MAL-I tolerates substitutions at position 3´, sialylation or sulphation even increasing the binding affinity (Geisler and Jarvis, 2011), thus explaining that no increase in the binding to Δ*siaB* was observed for this lectin, a behaviour also exhibited by WFA.

Besides the LOS, other Gal-containing structures on the bacterial surface could serve as lectin docking points. Indeed, the lectin-binding pattern to NTHi375Δ*lic2A*, lacking the proximal β-Gal in the Hep III ramification (Gal I, Fig. 4.1), is difficult to rationalise without considering recognition of other carbohydrate structures. Thus, although binding of VAA to Δ*lic2A* was decreased compared to the wild type strain, no significant alteration was observed in the case of RCA, MAL-I or WFA.

An intricate lectin-binding pattern was also observed for NTHi375Δ*lgtF*, lacking the glycosyltransferase responsible for adding Glc to Hep I. Here, binding of VAA and SBA was enhanced, what could be tentatively explained by an increase in accessibility of the Hep III extension carrying Gal, a trait also observed for the anti-NTHi serum (Fig. 4.2). However, recognition by the other Gal-specific lectins (RCA, MAL-I and WFA) was decreased. Of note, Δ*lgtF* was found to exhibit an altered behaviour compared to the wild type strain in terms of resistance to antimicrobial peptides, biofilm

Table 4.1. Binding of selected reference lectins to NTHi375 mutant strains[a]

	NTHi375 mutants						
Lectin	**Δ*siaB***	**Δ*lgtF***	**Δ*lic2A***	**Δ*lpsA***	**Δ*lgtF* Δ*lpsA***	**Δ*opsX***	**Δ*ompP5***
VAA	156 ± 5	165 ± 6	66 ± 6	38 ± 3	36 ± 1	31 ± 1	257 ± 10
RCA	140 ± 15	70 ± 4	105 ± 3	64 ± 5	32 ± 5	75 ± 15	64 ± 7
MAL-I	92 ± 5	80 ± 6	114 ± 8	55 ± 5	70 ± 4	61 ± 4	226 ± 33
SBA	140 ± 10	250 ± 25	132 ± 7	70 ± 8	76 ± 8	52 ± 3	283 ± 7
WFA	102 ± 5	82 ± 5	110 ± 5	92 ± 4	72 ± 4	44 ± 4	108 ± 8
WGA	47 ± 3	28 ± 3	58 ± 9	42 ± 5	22 ± 3	21 ± 3	59 ± 13
PSqL	80 ± 10	70 ± 10	80 ± 5	36 ± 3	41 ± 3	30 ± 4	72 ± 4
SNA	61 ± 4	45 ± 3	100 ± 6	85 ± 10	54 ± 9	52 ± 3	140 ± 10
ConA	170 ± 15	117 ± 7	140 ± 10	95 ± 4	82 ± 7	106 ± 6	272 ± 37

[a]Percentage of binding taking for each lectin the binding to wild type NTHi375 as 100%. Data shown corresponds to mean values calculated from binding intensities of triplicate samples of at least two different experiments to bacteria dilutions at OD_{600} of 1 and 0.6. Standard deviations of the mean are given.

growth, and epithelial and lung infection (Morey *et al.*, 2013), evidencing that this apparently small LOS truncation results in key phenotypic consequences. The reduced biofilm formation by Δ*lgtF* was particularly unexpected, as sialic acid residues, which are supposed to be present in the Hep III extension of Δ*lgtF,* are known to be important for biofilm growth (Swords *et al.*, 2004; Jurcisek *et al.*, 2005). Interestingly, Δ*lgtF* exhibited a significantly decreased binding of the Neu5Ac-specific lectins WGA, PsqL and SNA, indicating a lower availability of sialic acid residues for lectin recognition. Possible explanations could be a major change in the three-dimensional presentation of the LOS upon removal of PCho-Glc at Hep I, or a decrease in the sialic acid content of the LOS molecule due to dynamic compensatory changes in response to the truncation.

The binding pattern of ConA was also complex, with enhanced recognition of Δ*siaB*, Δ*lic2A* and Δ*lgtF* mutants, and similar or slightly decreased binding to the mutants with the shortest LOS versions (Δ*lpsA*, Δ*lgtF*Δ*lpsA* and Δ*opsX*). Increased ConA binding following desialylation was previously reported in histochemical studies (Takahashi, 1992). The PCho substitution at position 6 of the Glc unit at Hep I, as found in the wild type strain, might impede recognition of this Glc by ConA (Goldstein *et al.*, 1965). Indeed, Δ*lgtF* truncation seemed to favour recognition. In this mutant, Hep I is displayed as terminal moiety and may serve as possible docking point (Jaipuri *et al.*, 2008), therefore explaining the increased binding. Hep III, however, being acetylated and also substituted by PEtn, is not a likely target and, accordingly, binding to the Δ*lpsA* was comparable to that observed for the wild type strain. Furthermore, the small differences in binding to the Δ*lgtF*Δ*lpsA* and Δ*opsX* strains compared to wild type NTHi375 suggest that other surface carbohydrate structures, besides the LOS, must be recognised.

Overall, the noticeable effect of deletion of the outer core (Δ*lgtF*Δ*lpsA*) or all core sugars (Δ*opsX*) on the binding of most lectins supported the notion that the LOS could serve as lectin target. This idea was further nourished by the results obtained for the binding of the selected lectins to an isogenic NTHi375 mutant lacking P5, an outer membrane immunogenic protein that favours adhesion to mucosae (Miyamoto and Bakaletz, 1996) or mucin structures (Reddy *et al.*, 1996). In accordance, the Δ*ompP5* NTHi375 mutant was found to present impaired adhesion to epithelial cells and subsequent invasion (Euba *et al.*, 2015). Previous blotting analysis with the anti-PCho monoclonal antibody TEPC-15 had revealed more intense marks for Δ*ompP5* over wild

type NTHi375 colonies (personal communication of Dr. Junkal Garmendia), hinting at a higher PCho content in this mutant and, therefore, LOS expression, as PCho is a constituent of NTHi375 LOS (Bouchet *et al.*, 2003). And indeed, the amount of LOS isolated from Δ*ompP5* was significantly higher than that obtained from wild type NTHi375 (see section 4.1.5.1). As illustrated in Fig. 4.4 and detailed in Table 4.1, a moderate drop in the binding was observed for RCA, WGA and PSqL upon disruption of the *ompP5* gene, and no significant effect on WFA was found. However, the binding of SNA, and to a larger extent of VAA, MAL-I, SBA, and ConA to the Δ*ompP5* mutant was substantially increased compared to wild type NTHi375.

4.1.4 QCM bacteria-based biosensors: monitoring lectin binding kinetics directly on the bacterial surface

To confirm lectin recognition and strain-specific binding trends using a complementary approach, a novel application of the QCM technology was developed. RCA, WGA and ConA were selected as representative lectins with Gal, sialic acid/GlcNAc, and Man/Glc binding specificities, respectively. Wild type NTHi375 and the Δ*lgtF*Δ*lpsA* mutant were selected as reference strains, the latter showing in general the weakest binding signals for the tested lectins (Table 4.1). This new QCM approach was expected to provide additional insights into the interactions occurring at the bacterial surface, including information on binding kinetics, as successfully done previously for lectin binding to the eukaryotic cell surface (Pei *et al.*, 2012; Peiris *et al.*, 2012; Li *et al.*, 2013; Li *et al.*, 2015). The concept and development of QCM bacteria-based chips is described in detail in section 3.6.1. Based on the fairly comparable binding of ConA to NTHi375 and the Δ*lgtF*Δ*lpsA* mutant in the microarray assays (Table 4.1), this lectin was selected as bacteria capturer.

ConA covalently immobilised on activated LNB-carboxyl QCM chips via amine coupling was previously found to be operative for *ex situ* capture of cancer cells through recognition of cell surface carbohydrates (Li *et al.*, 2013), supporting the applicability of our setup. Bacteria cells were thus captured onto the surface of ConA-derivatised QCM chips and fixed in formaldehyde, so that they become an integrated part of the sensor chip. Fixation is key to allow the use of efficient regeneration solutions for completely removing bound lectins between successive binding assays. Regeneration of non-fixed mammalian cell chips following lectin-binding assays using a high

concentration of a competitive sugar had been reported (Li *et al.*, 2013; Li *et al.*, 2015). Still, the use of this cell chip was limited to a few injection cycles. Furthermore, cell outflow might occur if the sugar can also compete with the capturing-lectin–cell interaction. In contrast, the formaldehyde-fixed ConA-bacteria chips here developed could be completely regenerated using a pulse injection of 10 mM glycine, pH 1.2, 0.5 M NaCl, were stable for more than 50 injection cycles, and could be stored at 4 ºC for further use (Kalograiaki *et al.*, 2016).

Bacteria coverage of the chip surface before and after the binding assays (Fig. 4.5, *a* and *b*) was assessed by DAPI staining of nucleic acids and fluorescence microscopy, showing a widespread capture of cells over the entire surface and proving retention of the bacteria after injection and regeneration cycles. Furthermore, lectin binding to the immobilized bacteria was visualised by fluorescence microscopy following *ex situ* incubation with FITC-ConA, evidencing the availability of ligands for ConA on the bacterial surface all over the chip (Fig. 4.5, *c*).

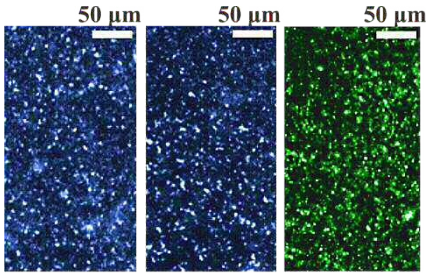


Fig. 4.5 Fluorescence microscopy evaluation of the wild type 375 chip. Bacteria coverage of the chip surface was evaluated before *(a)* and after *(b)* 30 lectin-binding and regeneration cycles by *ex situ* DAPI staining and visualization using fluorescence microscopy; *(c)* FITC-ConA binding to the bacteria chips after the lectin binding assays.

Having confirmed the efficiency of ConA-derivatised chips for bacteria capturing, the stability of the formaldehyde-fixed ConA-bacteria chips and their applicability for lectin binding studies, we examined the interaction of RCA, WGA and ConA with NTHi375 and Δl*gtF*Δ*lpsA* chips. After stabilisation of the resonant frequency under continuous flow of running buffer, lectins were injected in a range of increasing concentrations. Examples of the recorded sensograms are shown in Fig. 4.6.

Strain- and lectin-specific responses were clearly observed. A first parameter deserving attention was the maximum frequency shift induced by injection of each lectin, $-\Delta f$ (Table 4.2). The amplitude of the response was proportional to the amount of lectin bound to the chip at the end of the injection, as well as to the molecular mass of the respective lectin. Since all three lectins were tested on the same surfaces (i.e., against

the same amount of bacteria on the chip), the frequency shift can be used to compare lectin binding to each chip after normalisation for the lectin molecular mass.

As detailed in Table 4.2, WGA and especially RCA bound to the NTHi375 chips to a significantly greater extent than ConA. Of note, the observed trend was analogous to that found for the hapten-inhibitable binding in the microarray assays (Fig. 4.3, lower panel). Interestingly, ConA binding to the QCM chips was completely abolished when

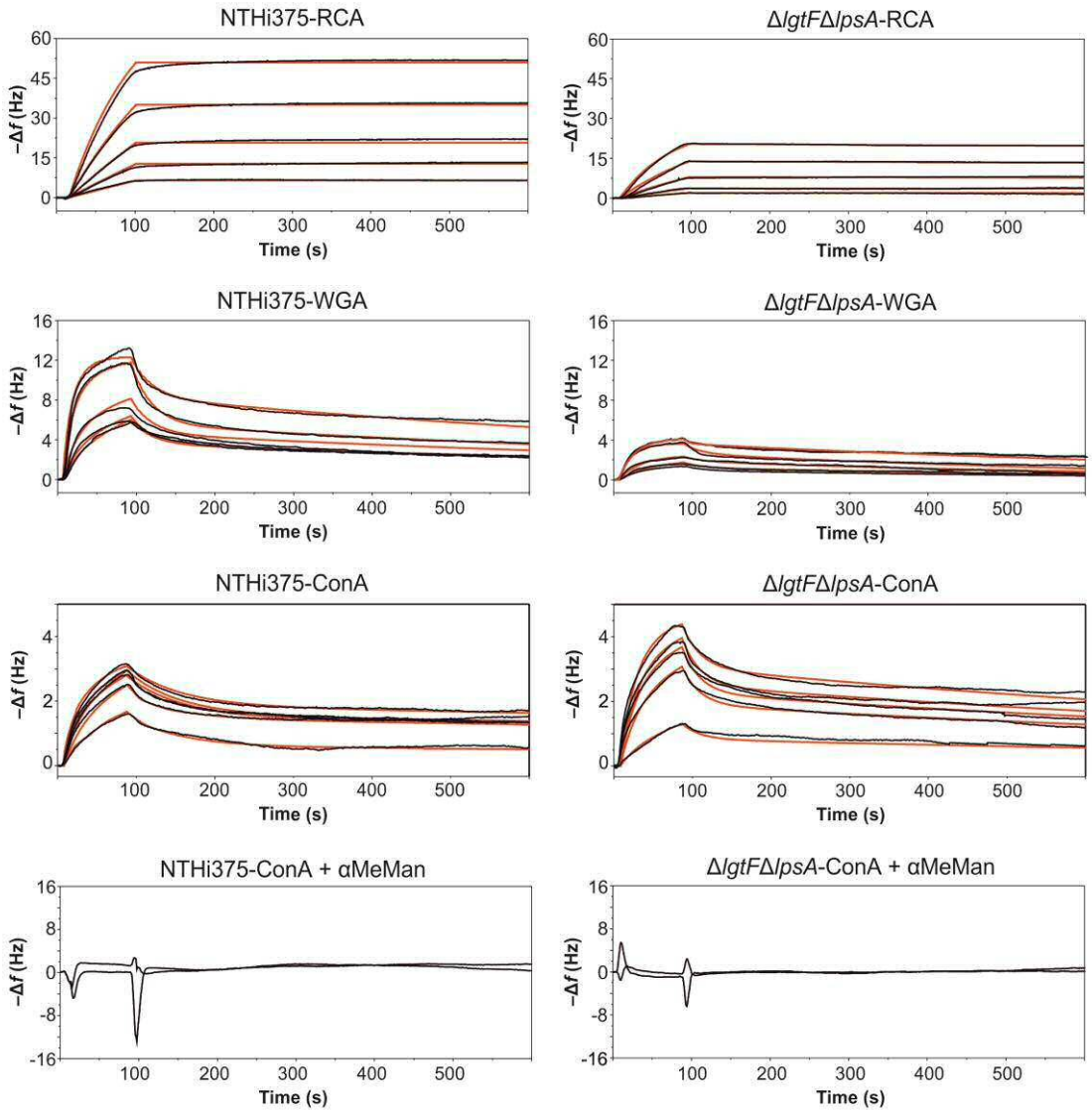


Fig. 4.6 Representative sensograms illustrating the binding of RCA, WGA, and ConA to NTHi375 and Δ*lgtF*Δ*lpsA*. Lectins were injected at the following concentrations: 0.3, 0.6, 1.25, 2.5 and 5 µg/ml for RCA; 6.25, 12.5, 25, 50 and 100 µg/ml for WGA; and 10, 60, 90, 100 and 150 µg/ml for ConA. The black lines in the sensogram represent experimental data, and the red lines the fits obtained using 1:1 (RCA) or 1:2 (ConA and WGA) models. In the competition experiments (bottom panels), the binding of ConA at concentrations of 100 and 150 µg/ml was examined after pre-incubation of the lectin with 0.2 M α-Me-Man for 1 h before injection.

the lectin was pre-incubated with 0.2 M α-Me-Man before injection (Fig. 4.6), proving that the monitored interactions were carbohydrate-dependent. Binding of ConA to the NTHi375 arrays, however, was only inhibited to a small extent in the presence of this hapten, suggesting that after the primary carbohydrate-mediated interaction detected in the dynamic QCM analysis, further interactions are established at the longer contact times of the static microarray binding assays.

Compared to NTHi375, the binding of RCA and WGA to Δ*lgtF*Δ*lpsA* chips was noticeably smaller (41% and 33%, respectively, of the binding to the wild type strain), whereas for ConA a small increase (33%) in the $-\Delta f$ values was observed (Table 4.2). Therefore, the lower $-\Delta f$ induced by injection of RCA and WGA over the mutant chips should be ascribed to a different availability on the bacterial surface of carbohydrate structures recognised by each lectin rather than to a lower bacteria content. What is more, the observed trend was again analogous to that found in the microarray binding

Table 4.2 QCM analysis of the binding of RCA, WGA and ConA to NTHi375 wild type and Δ*lgtF*Δ*lpsA* strains

Parameter	RCA		WGA		ConA	
	NTHi375	Δ*lgtF*Δ*lpsA*	NTHi375	Δ*lgtF*Δ*lpsA*	NTHi375	Δ*lgtF*Δ*lpsA*
$-\Delta f$ (Hz)	47 ± 1.5	19.5 ± 1.5	12 ± 1	4.01 ± 0.02	3.14 ± 0.02	4.45 ± 0.07
Normalised $-\Delta f$ (Hz kDa^{-1})	0.39	0.16	0.33	0.11	0.03	0.04
k_{a1} (M^{-1} s^{-1})	1.75 ± 0.01 × 10^{5}	1.19 ± 0.01 × 10^{5}	2.15 ± 0.06 × 10^{4}	4.2 ± 0.1 × 10^{4}	2.01 ± 0.05 × 10^{4}	2.8 ± 0.4 × 10^{4}
k_{d1} (s^{-1})	8.76 ± 0.02 × 10^{-7}	6.29 ± 0.05 × 10^{-5}	8.99 ± 0.01 × 10^{-4}	1.63 ± 0.01 × 10^{-3}	2.2 ± 0.1 × 10^{-4}	3.4 ± 0.1 × 10^{-3}
K_{D1} (M)	5.0 ± 0.1 × 10^{-12}	5.29 ± 0.01 × 10^{-10}	4.2 ± 0.1 × 10^{-8}	3.88 ± 0.01 × 10^{-8}	1.1 ± 0.7 × 10^{-8}	1.2 ± 0.2 × 10^{-7}
k_{a2} (M^{-1} s^{-1})	—	—	4.6 ± 0.1 × 10^{4}	1.5 ± 0.15 × 10^{4}	5.25 ± 0.4 × 10^{4}	4.6 ± 0.2 × 10^{4}
k_{d2} (s^{-1})	—	—	3.88 ± 0.01 × 10^{-2}	4.62 ± 0.01 × 10^{-2}	2.1 ± 0.1 × 10^{-2}	4.7 ± 0.15 × 10^{-2}
K_{D2} (M)	—	—	8.4 ± 0.2 × 10^{-7}	3.1 ± 0.3 × 10^{-6}	4.0 ± 0.01 × 10^{-7}	1.0 ± 0.1 × 10^{-6}

The association rate (k_a), dissociation rate (k_d) and dissociation constant (K_D) are derived from the fitting of the experimental data to 1:1 (RCA) or 1:2 (WGA and ConA) models. Frequency shifts were recorded upon injection of 5 μg/ml of RCA, 100 μg/ml of WGA or 150 μg/ml of ConA. Molar concentrations were calculated considering molecular masses of 120 kDa for RCA, 36 kDa for WGA, and 104 kDa for ConA. Normalised $-\Delta f$ values were calculated by dividing $-\Delta f$ by the molecular mass of each lectin. Standard deviations of mean values are given.

assays, with ConA being the less affected by the outer core truncation and WGA the most (Fig. 4.6 and Table 4.2). At any rate, frequency shifts achieved were higher than 0.9 Hz (also see Fig. 4.6, WGA), defined as the lowest $-\Delta f$ detection limit for Attana QCM-chips bearing crystals of 10 MHz-nominal frequency and circular electrodes (Anderson *et al.*, 2007). This permitted the safe calculation of kinetic parameters (k_a and k_d).

For calculation of the kinetic parameters (k_a and k_d), the experimental data were fitted using 1:1 or 1:2 binding models. The first model was found suitable for fitting the binding profiles of RCA (Fig. 4.6 and Table 4.2). Its binding kinetics to the wild type strain was characterised by a relatively fast association rate (in the range of 10^5 M^{-1} s^{-1}) and an extremely slow dissociation rate, in the order of 10^{-7} s^{-1}. This k_d indicates the formation of an extremely stable complex that does not dissociate over the time frame of the experiment, leading to a very high binding affinity, in the picomolar range. Worth to be mentioned, dissociation rates in the range of 10^{-7} s^{-1} or slower require extremely long measurement times to observe a significant shift in the dissociation phase, making these strong binding events difficult to quantitate. In comparison, the association rate to Δ*lgtF*Δ*lpsA* (Table 4.2) was similar, but dissociation was 100-fold faster, resulting in a proportionally lower binding affinity.

On the other hand, a 1:2 model was required for fitting the experimental data obtained for WGA and ConA (Fig. 4.6), what could indicate different accessibilities of the bound target due to specific micro-environments, uneven clustering of the target on the bacterial surface, locally favouring multivalent interactions, or even binding to different targets with dissimilar binding affinities. The rather broad nominal binding-specificity of these two lectins, which are able to recognise different monosaccharides (sialic acid and GlcNAc for WGA; Glc, Man and manno-Hep for ConA) makes the existence of multiple targets highly likely.

The difference in affinity between the two sets of binding sites on the surface of both NTHi375 and Δ*lgtF*Δ*lpsA* primarily arose from a faster dissociation rate from the low-affinity sites, with k_d values on the order of 10^{-2} s^{-1}, whereas the k_a values for both low- and high-affinity sites were in the range of 10^4 M^{-1} s^{-1} (Table 4.2). However dissociation rates for the high-affinity sites were faster for Δ*lgtF*Δ*lpsA* compared to the wild type strain, resulting in a lower binding affinity of both WGA and ConA towards the mutant.

The K_Ds calculated from the kinetic parameters were around 10^{-6} M for the low-affinity sites of WGA and ConA, and ranged from 10^{-7} to 10^{-8} M for the high-affinity sites, while for RCA K_D values in the order of 10^{-10}-10^{-12} M were obtained. Although these affinities are noticeably higher than those exhibited by these lectins when binding to purified glycan structures, they are comparable to those reported for the binding to cell surfaces. As example, K_D values in the 10^{-7}-10^{-9} M range were determined by QCM for the binding of RCA, WGA, ConA and other lectins to a diversity of cancer cells (Peiris *et al.*, 2012; Li *et al.*, 2013; Li *et al.*, 2015). It seems reasonable to presume that the complex mixture and density of epitopes displayed on the cell surfaces may favour the establishment of high-affinity interactions. In contrast, K_D values in the 10^{-4} to 10^{-5} M range were also determined for the binding of HPA and *Ulex europaeus* agglutinin I to certain cancer cells (Peiris *et al.*, 2012), evidencing that high-affinity interactions are lectin- and cell-specific. Similarly, the binding affinities here determined for the binding of RCA, WGA, and ConA to NTHi375 and Δ*lgtF*Δ*lpsA* were lectin- and strain-specific, the three lectins exhibiting a higher affinity for the wild type strain than for the mutant (Table 4.2).

Overall, the results of the QCM analysis fully agreed with the binding trends observed in the microarray analysis, supporting the utility of the microarray-based approach to assess accessibility for lectin recognition of carbohydrate structures on the surface of NTHi.

4.1.5 Evaluation of lectin binding to the isolated NTHi375 LOS

Prompted by the results of the bacteria-based microarray and QCM analyses, hinting at the LOS as likely target for many of the tested lectins, we examined the binding of the lectins to the isolated LOS using microarray and NMR approaches.

4.1.5.1 LOS quantitation and validation of LOS microarrays

As mentioned before (section 4.1.3), NTHi375Δ*ompP5* was suspected to overexpress LOS based on a higher reactivity of this mutant towards the anti-PCho antibody TEPC-15. Therefore, the LOSs of NTHi375 wild type and Δ*ompP5* mutant strains were extracted (described in section 3.2) and quantified by a combination of the Purpald assay (Fig. 4.7, *a*) and densitometry of LOS bands upon DOC-PAGE electrophoresis

and silver staining (Fig. 4.7, *b*), following an optimised protocol. *S. minnesota* smooth and Rb LOSs were used as reference.

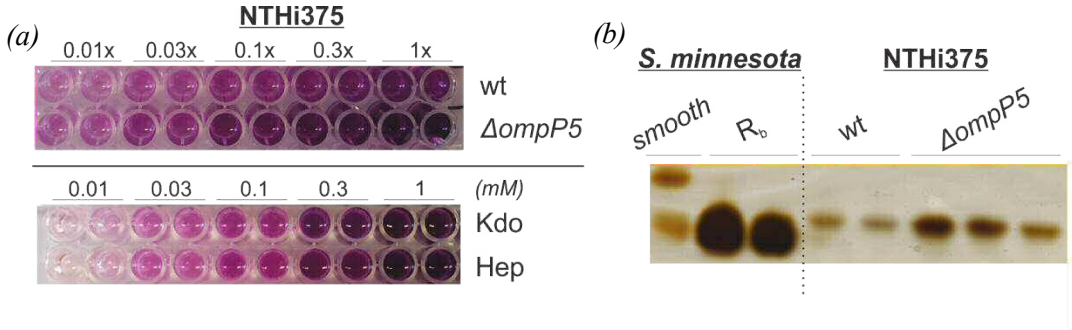


Fig. 4.7 Quantitation of the LOS isolated from NTHi375 wild type and Δ*ompP5* mutant strains. *(a)* Colorimetric quantitation by the Purpald assay. LOS dilutions from the stock solution are indicated (in duplicates). *(b)* Electrophoretic profiles of the isolated LOSs; Sample loaded: (i) 1 µg, (only shown the two lower M_w bands); (ii) 1 µg, single band; (iii) 0.7 µg; (iv) 2 µl stock; (v) 1 µl stock; (vi) 2 µl (dil 1:10); (vii) 2 µl (dil 1:20); (viii) 2 µl (dil 1:30).

Although batch-to-batch variations in the relative yields of the isolated LOS were observed, the estimated quantity of LOS extracted from the Δ*ompP5* mutant was always significantly higher than that extracted from a wild type NTHi375 culture of equal biomass. As Δ*ompP5* is an isogenic mutant of NTHi375 that keeps unaltered the enzymatic machinery involved in LOS biosynthesis, their LOSs are expected to be structurally identical. Consequently, the higher yield in LOS extraction from Δ*ompP5* cannot derive from a differential partition index in the extraction media due to structural divergences, directly pointing to a higher LOS expression in this mutant compared to wild type NTHi375. Thus, apparently, deletion of this major outer membrane protein results in LOS overexpression, probably due to compensatory changes in the bacterial surface.

LOSs extracted from wild type and Δ*ompP5* NTHi375 were printed in manual microarrays along the *S. minnesota* Rb LOS as control, and the utility of the arrays for detecting recognition by LOS-

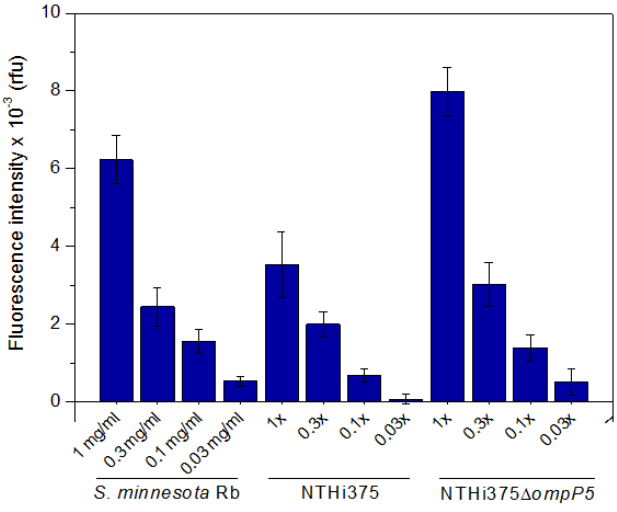


Fig. 4.8 Anti-lipid A binding to manually printed LOS microarrays. Each LOS was printed in decreasing concentrations as sextupliclates. Concentrations refer to the dilution of the stock solution.

binding molecules was examined by testing the binding of a polyclonal antibody against *E. coli* O157-lipid A (Lee *et al.*, 2004), of common structure with NTHi-lipid A (Helander *et al.*, 1988; Risberg *et al.*, 1999; Garmendia *et al.*, 2014). *Salmonella*-lipid A also shares structural features (Qureshi *et al.*, 1985) with NTHi-lipid A – both contain (14:0) myristate chains (frequently hydroxylated at position 3') and are phosphorylated at position 4' of the non-reducing GlcN.

As shown in Fig. 4.8, arrayed NTHi LOSs were recognized efficiently by the antibody. Moreover, fluorescence signals observed for the Δ*ompP5* LOS were stronger, in line with the quantitation results (Fig. 4.7).

4.1.5.2 Lectin binding to LOS microarrays

As already mentioned, the LOS of Δ*ompP5* should be structurally identical to that of the wild type strain. Therefore, since the amount of LOS isolated from wild type NTHi375 was very limited, the Δ*ompP5* LOS was used to investigate the binding of the panel of selected lectins. Microarray data (Fig. 4.9) indicated direct LOS binding of weak to very strong intensity for WGA, ConA, and VAA. In all cases, the binding was carbohydrate-mediated, as it was inhibited in the presence of the appropriate hapten. The weak binding observed for WGA could be attributed to the recognition of the minor sialylated LOS glycoforms, what would be supported by the significant drop in WGA-binding to the entire bacteria upon disruption of the *siaB* gene. The N-acyl GlcN residues of lipid A, which are expected to be more accessible in the isolated LOS than on the intact bacterial surface, might also serve as docking points. The binding of ConA indicates that Glc and/or manno-Hep residues within the LOS can be recognized by this lectin. Regarding VAA, recognition of the terminal α-Gal moiety, present in the major LOS glycoform as part of a globotriose in the Hep III extension, could account for the strong binding of this lectin.

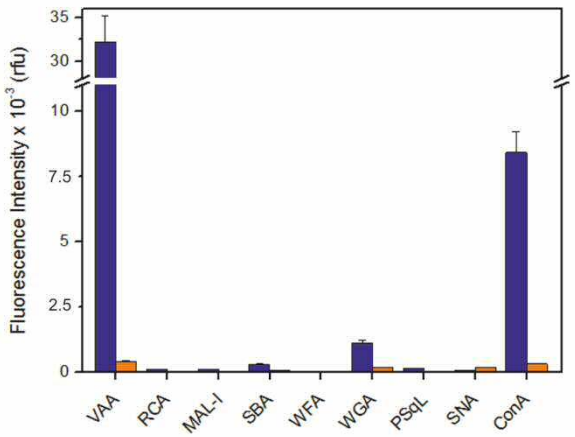


Fig. 4.9 Recognition of Δ*ompP5* LOS by reference lectins. The LOS was printed at a concentration of 1 mg/ml and the binding of the lectins was tested in the absence (blue) or presence (orange) of their competitive hapten.

Of note, no binding of RCA to the LOS microarrays was detected. Indeed, although VAA and RCA exhibited noticeable binding to NTHi375, evidencing the availability of galactose-containing structures on the bacterial surface, LOS truncation had disparate consequences on the binding of the two lectins (Fig. 4.4). In particular, binding of VAA to the Δ*lic2A* mutant, lacking Gal in the LOS, was importantly decreased compared to the wild type strain, while this truncation had no significant effect on RCA binding. Furthermore, the binding of VAA to the Δ*ompP5* mutant, here found to overexpress the LOS, was substantially increased compared to the wild type strain, whereas a drop in the binding was observed for RCA (Fig. 4.4 and Table 4.1). Altogether, the results strongly indicate that VAA and RCA recognise different ligands on the bacterial surface. At this point it is important to mention that both VAA and RCA are galactose-specific $(AB)_2$-type agglutinins that show a high structural homology but also exhibit different fine-oligosaccharide binding specificities. In particular, VAA is able to bind both α- and β-galactosides (Lee *et al.*, 1992), whereas RCA shows a clear preference for terminal β-Gal (Podder *et al.*, 1974).

The binding of ConA to the LOS microarrays was also fully consistent with the increased binding of this lectin to the Δ*ompP5* mutant compared to wild type NTHi375. However, for the rest of lectins, there was no correlation between their binding to LOS microarrays and to the entire Δ*ompP5* bacteria. First, although WGA was capable to bind weakly to the isolated LOS, its binding to Δo*mpP5* was smaller than to wild type NTHi375, what could be due to a reduced accessibility to the recognized epitope resulting from LOS overexpression. On the other hand, although the binding of MAL-I, SBA, and SNA to the Δ*ompP5* mutant was substantially increased compared to wild type NTHi375, no binding to the LOS microarrays was detected for these lectins, suggesting that they could recognise other carbohydrate structures on the bacterial surface. Still, the possibility that the specific presentation of the LOS upon immobilisation in the microarrays might substantially differ from that on the entire bacteria and preclude lectin recognition cannot be excluded. Overall, the results stressed the importance of examining recognition directly at the bacterial surface, where the native environment, density and accessibility of the recognised epitopes are preserved.

4.1.6 Comparative analysis of the binding of VAA and RCA to NTHi Rd KW20 and its transformed Rd *hmw1*$_{strain12}$ strain

Altogether, the results above reported strongly indicate that the major ligand(s) recognized by VAA and RCA on the bacterial surface are different. Besides the LOS, other Gal-containing structures exposed on the surface of NTHi could serve as lectin docking points. A relevant candidate is the high molecular weight glycoprotein HMW1, which is glycosylated at multiple asparagine residues with N-linked simple glucose and galactose moieties or with disaccharide units, the latter only present at sites with glucose first linked to asparagine (Grass *et al.*, 2010). In order to evaluate HMW1 as possible docking site for VAA and/or RCA, we extended our study to a capsule-deficient NTHi laboratory strain, Rd KW20 (hereafter to as Rd), which lacks the *hmw1* operon (and its paralog *hmw2*). Direct comparison of lectin binding to NTHi375 and Rd strains is not appropriate, since the two strains have a distinct genetic background (Hood *et al.*, 2001; Cody *et al.*, 2003). We thus took advantage of the availability of Rd *hmw1*$_{strain12}$, a transformed Rd strain carrying the *hmw1* operon of NTHi strain 12 (also known as R2846). This was the prototypic HMW-positive strain in which the glycoprotein was originally identified, isolated and sequenced (Barenkamp and Leininger, 1992), and most extensively characterised. The HMW1$_{strain12}$ adhesin shares a homology of 93% with that of NTHi375. The binding of VAA and RCA to Rd and Rd *hmw1*$_{strain12}$ was comparatively examined, starting with the bacteria microarray analysis (Fig. 4.10).

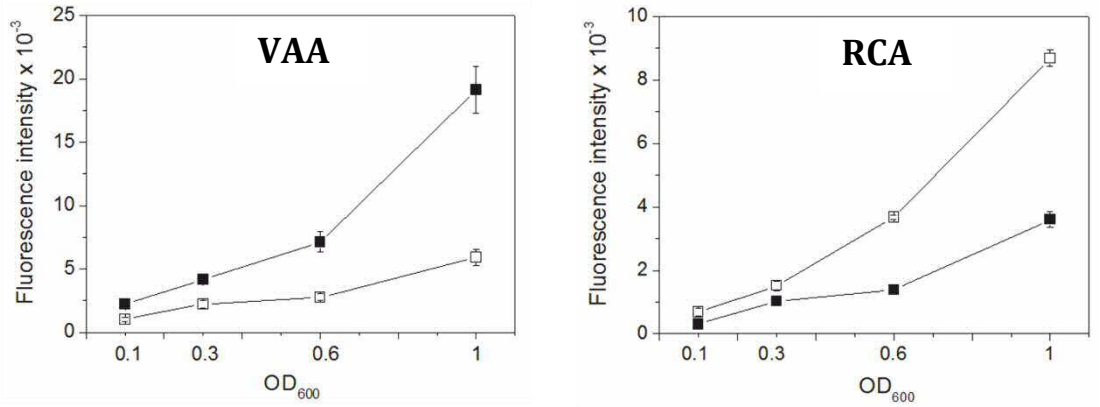


Fig. 4.10 Binding of VAA and RCA to Rd and Rd *hmw1*$_{strain12}$. Different dilutions of Rd (□) and Rd *hmw1*$_{strain12}$ (■) were immobilised on NC-coated microarray slides, and the binding of the two lectins was tested as described before. Normalised fluorescence intensity values are expressed in rfu. Data shown correspond to the mean of at least four independent experiments. Error bars indicate the standard deviation of the mean.

Binding of VAA to the transformed Rd *hmw1*$_{strain12}$ strain was noticeably higher than to Rd, what indicates that the HMW1 glycoprotein may also serve as ligand for this lectin (Fig. 4.10). In striking contrast, the binding of RCA was significantly lower, suggesting that the Gal moieties "decorating" HMW1$_{strain12}$ are not docking sites for RCA and, what is more, that the expression of HMW1 at the bacterial surface even decreases the availability of ligands for RCA recognition.

Following our combined microarray and QCM approach for exploring lectin recognition on the bacterial surface, we next performed binding assays of VAA and RCA to Rd and Rd*hmw1*$_{strain12}$ QCM chips. In all cases, data could be fitted to a one-class-of-sites model. The parameters obtained for the binding of RCA to Rd and Rd*hmw1*$_{strain12}$ were similar (Table 4.3). As observed for NTHi375 (Table 4.2), the RCA binding kinetics were characterised by relatively fast association rates and extremely slow dissociation rates, leading to binding affinities in the picomolar range. For VAA, the parameters obtained for the binding to both strains were also similar (Table 4.3), with association rates comparable to those observed for RCA. However, dissociation rates were significantly faster for VAA, resulting in a lower binding affinity towards both Rd and Rd*hmw1*$_{strain12}$.

Table 4.3 QCM analysis of the binding of VAA and RCA to Rd and Rd *hmw1*$_{strain12}$

Parameter	VAA		RCA	
	Rd	Rd *hmw1*$_{strain12}$	Rd	Rd *hmw1*$_{strain12}$
k_a (M^{-1} s^{-1})	$6.71 \pm 0.01 \times 10^5$	$4.42 \pm 0.01 \times 10^5$	$6.47 \pm 0.01 \times 10^5$	$2.7 \pm 0.1 \times 10^5$
k_d (s^{-1})	$4.37 \pm 0.02 \times 10^{-4}$	$3.39 \pm 0.07 \times 10^{-4}$	$9.55 \pm 0.03 \times 10^{-7}$	$2.9 \pm 0.5 \times 10^{-7}$
K_D (M)	$6.5 \pm 0.03 \times 10^{-10}$	$7.67 \pm 0.17 \times 10^{-10}$	$1.48 \pm 0.01 \times 10^{-12}$	$1.02 \pm 0.01 \times 10^{-12}$

[a]Kinetic values are derived from fitting of the experimental data to 1:1 model for both lectins. Frequency shifts were recorded upon injection of 2.5 µg/ml of VAA or RCA. Molar concentrations were calculated considering molecular masses of 114 kDa for VAA and 120 kDa for RCA. Standard deviations of mean values are given.

Noteworthy, the LOS of the Rd strain (Fig. 4.11) is structurally very similar to that of NTHi375 (Fig. 4.1). Its major glycoform, representing nearly 60% of LOS forms, presents a terminal Lac moiety anchored to Hep III, whereas two additional glycoforms bear a terminal αGal(1,4)-linked to Lac, as found in the NTHi375 LOS, or a further elongation by addition of β(1,3)-linked GalNAc (Fig. 4.11) (Risberg *et al.*, 1999; Hood

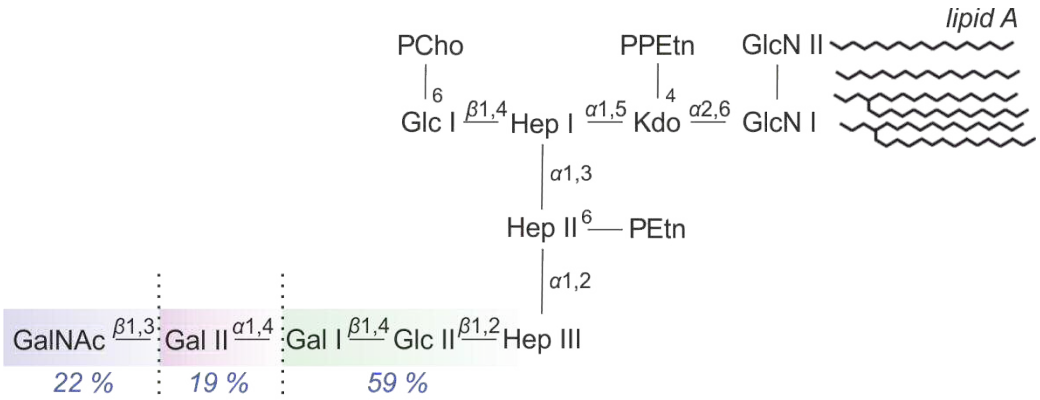


Fig. 4.11 LOS structure proposed for NTHi Rd strain. The major glycoform (Hex3) bears the Hep inner core and three Hex*p* moieties with Gal I at the non-reducing end, representing 59% of total species. Minor glycoforms (Hex4 & Hex5) were also detected by ESI/MS and their occurrence was assigned as noted (Risberg *et al.*, 1999). Abbreviations generally stand as for Fig. 4.1 for coherence, while GalNAc refers to galactosamine.

et al., 2001). Thus, the Rd LOS could constitute a ligand for both VAA and RCA. Therefore, considering the disparate binding behaviour of VAA and RCA to the isolated NTHi375 LOS, as well as the lower binding affinity of VAA towards the Rd strains compared to RCA, evidenced by the QCM analysis, a direct comparison of the binding of the two lectins to the Rd LOS appeared essential.

To this aim, the LOS of strain Rd was extracted and quantified, Rd LOS microrrays were prepared as described for NTHi375 LOS (section 4.1.5.1), and binding assays with RCA and VAA were carried out. Strong and lactose-inhibitable binding of RCA to the Rd LOS was observed (Fig. 4.12), proving that the terminal Lac moiety present in the major Rd LOS glycoform can serve as docking point for this lectin. In striking contrast, only weak binding was observed for VAA. This result was particularly unexpected, as the major glycoform of the NTHi375 LOS, to which VAA was found to bind very strongly, represents ~19% of the Rd LOS. Hence, in order to confirm these results and also to get further insights into the NTHi375/Rd LOS epitopes recognized by RCA and VAA, a saturation transfer difference (STD) NMR analysis of LOS binding by the two lectins was carried out.

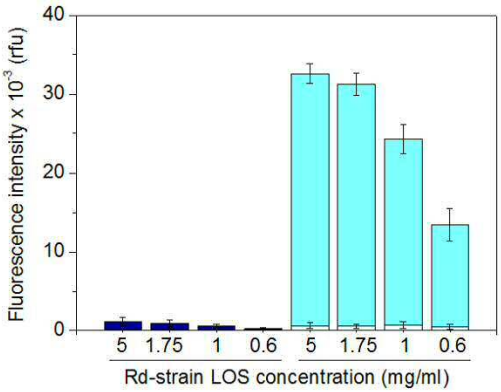


Fig. 4.12 Binding of VAA and RCA to microarray-printed Rd LOS. The binding of 74 µg/ml VAA (blue) and 8 µg/ml RCA (cyan) was tested in the absence or presence of 100 mM lactose (white). Binding intensities correspond to the mean of triplicate data and the error bars indicate the standard deviation.

4.1.7 STD NMR analysis of the binding of VAA and RCA to the NTHi375Δ*ompP5* and Rd LOSs

STD NMR is a sensitive method to study weak to medium interactions (K_D in the μM range), like those frequently found between glycans and lectins. Moreover, it is very accurate for identifying the atoms of a ligand in close contact with the receptor in a process called "ligand epitope mapping" (Mayer and Meyer, 2001). To this aim, assignment of the NMR resonances of the ligand is required. Therefore, in order to perform an STD NMR analysis of the binding of VAA and RCA to the LOS isolated from strains NTHi375Δ*ompP5* and Rd, we first approached the structural elucidation of these LOSs.

4.1.7.1 NMR structural characterisation of NTHi375Δ*ompP5* and Rd LOSs.

The ^{1}H-NMR spectra of the isolated LOSs in D_2O showed only broad undefined signals possibly due to lipid-mediated aggregation or even micelle formation, as previously reported for Rd LOS (Risberg *et al.*, 1999). In order to overcome this problem, the LOS preparations were hydrolyzed with acetic acid under mild conditions affecting the acid-labile Kdo–lipid A linkage, thereby yielding insoluble lipid A and soluble core oligosaccharides (OSs) derived from the various LOS glycoforms. The hydrolysis resulted in considerable improvement in signal resolution and intensity, giving rise to

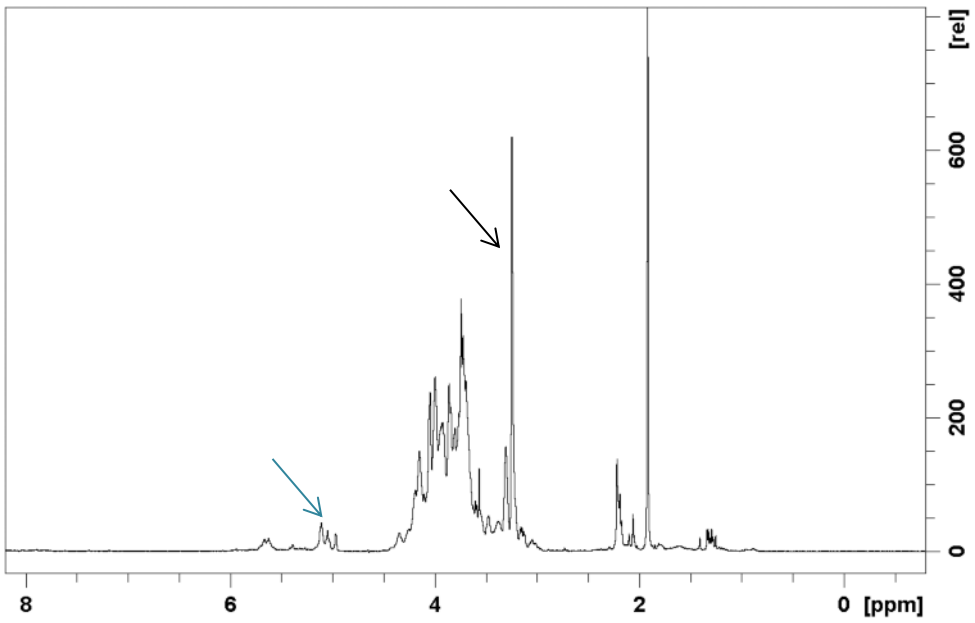


Fig. 4.13 ^{1}H-NMR spectrum of the OSs derived from NTHi375Δ*ompP5* LOS. Acquired in deuterated PBS at 310 K with water-signal suppression. The black arrow indicates signals corresponding to the PCho methyl protons. The blue arrow indicates those corresponding to the anomeric protons of *α*-sugars.

well-defined resonances particularly evident in the anomeric proton region (4.4-5.7 ppm) of the spectrum, as shown in Fig. 4.13 for Δ*ompP5* LOS. A notable feature of the spectra was the intense singlet at 3.25 ppm, attributed to the PCho methyl protons (Fig. 4.13).

A collection of ^{1}H-^{13}C HSQC, TOCSY and ROESY experiments (Figs. 4.14, 4.15 and Annex II) helped to the resolution of signals overlapping in the ^{1}H dimension, and enabled the assignment of chemical shifts and coupling constants of the anomeric protons (reported in Table 4.4 and Annex II).

Numerous resonances were visible in the Hep anomeric region of the spectra (Fig. 4.14, inset). Indeed, mild acid hydrolysis may result in partial cleavage of PEtn from the Hep moieties, elimination of the phosphate group from position 4 of Kdo (Carof *et al.*, 1987), and formation of several anhydro-forms of this sugar , altogether accounting for the heterogeneity of anomeric signals for the Hep moieties (Schweda *et al.*, 1993; Risberg *et al.*, 1999).

Focussing on NTHi375Δ*ompP5* OSs, for which detailed NMR data had not been reported so far, anomeric resonances were observed in the ^{1}H-NMR spectrum at δ =

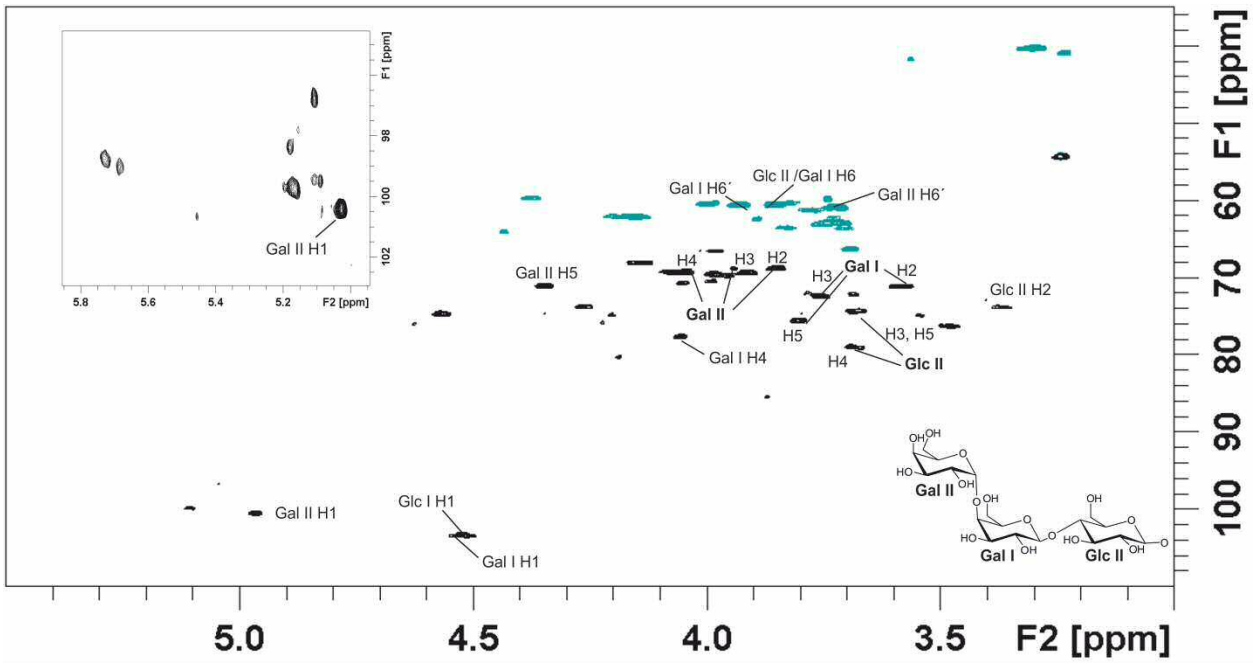


Fig. 4.14 Section of the multiplicity-edited ^{1}H-^{13}C HSQC spectrum of core oligosaccharides derived from NTHi375Δ*ompP5* LOS. Peak assignment is denoted for the terminal trisaccharide moiety (Hex4 glycoform) protons and the anomeric proton of Glc I. CH_2 are shown in petrol; CH and CH_3 in black. Chemical shifts and coupling constants of the anomeric protons are reported in Table 4.4. The insert corresponds to the *α*Gal and Hep anomeric proton region, δ from 5 to 6 ppm.

5.03-5.14 and 5.60-5.68 ppm (Fig. 4.13, Table 4.4) corresponding to the three Hep residues of the inner core. Coupling constants were not resolved for all of them, but their identities were further confirmed by the occurrence of ROE cross-peaks between the respective H1-H2 intra-residue resonance pairs (Table 4.5). In addition, intense transglycosidic ROE contacts between the proton pairs Hep III H1/Hep II H2 and Hep II H1/Hep I H3 (Table 4.5) were in agreement with the sequence of the inner-core trisaccharide being L-*α*-D-Hep-(1,2)-L-*α*-D-Hep-(1,3)-L-*α*-D-Hep-(1→), as previously described by NMR and mass spectroscopy for this and other Hi strains (Phillips *et al.*, 1992; Borrelli *et al.*, 1995; Masoud *et al.*, 1997; Risberg *et al.*, 1999; Fox *et al.*, 2006) .

Regarding Hex residues, the relatively large $J_{1,2}$-values of the anomeric resonances (Table 4.4) observed at δ = 4.52 (J = 8.2 Hz), 4.53 (J = 7.5 Hz) and 4.68 (J = 7.5 Hz) in the TOCSY spectrum, indicated that the corresponding residues have *β*-configuration. The anomeric signal at 4.47 ppm (J not resolved) was ascribed to the terminal *β*-linked galactosyl moiety of a Hex3 glycoform. A ROE cross-peak between H1 of Gal I and the overlapping signals of H3/H4/H5 of Glc II confirmed this notion (Fig. 4.15). In contrast, the anomeric signal at 4.97 ppm exhibited a small $J_{1,2}$ coupling (3.7 Hz), consistent with an *α*-linked saccharide. This signal was assigned to the terminal *α*-Gal*p* (Gal II) of a

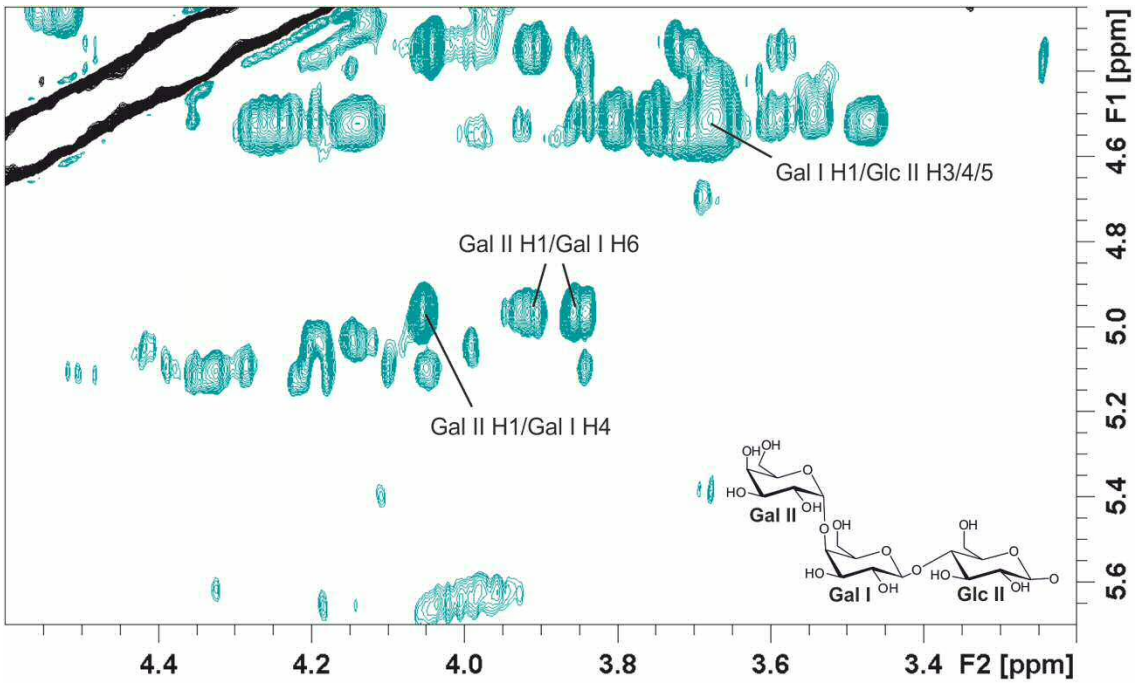


Fig. 4.15 Detail of ROESY spectrum of NTHi375Δ*ompP5*-derived OSs.
ROE connectivities for the globotriose component of the Hex4 glycoform. The spectrum was acquired with 300 ms spin lock mixing time.

Table 4.4 ^{1}H-^{13}C NMR chemical shifts for core oligosaccharides derived from OS glycoforms of NTHi375Δ*ompP5*

Isoform	Residue	Sugar unit	$^{3}J_{H,H}$ †	H1 (C1)	H2 (C2)	H3 (C3)	H4 (C4)	H5 (C5)	$H6_A$ (C6)	$H6_B$	$H7_A$ (C7)	$H7_B$
all	Hep I	3,4)-l-*α*-D-Hep*p*-(1	N/R	5.09-5.14 (99.5-99.9)	4.05-4.06 (69.2)	3.70 (78.9)	4.25* (78.1)	N/D	4.14* (64.5)	N/D	N/D	N/D
all	Hep II	2)-l-*α*-D-Hep*p*-(1	3.8	5.60-5.68 (99.2)	4.32 (78.8)	4.02 N/D	3.68 (72.8)	N/D	N/D	N/D	3.75 (65.0)	3.92 (61.2)
all	Hep III	2)-l-*α*-D-Hep*p*-(1	N/R	5.10-5.13 (99.7)	4.19; 4.21* (78.2)	4.02 N/D	N/D	N/D	N/D	N/D	N/D	N/D
all	Glc I	*β*-D-Glc*p*-(1	8.2	4.52 (103.3)	3.38 (73.8)	3.48 (76.2)	3.55 (74.8)	3.60* (74.9)	4.14 (64.5)	4.26 (64.5)		
all	Glc II	4)-*β*-D-Glc*p*-(1	7.5	4.68 (101.8)	3.38 (72.7)	3.69 (74.3)	3.69 (78.8)	3.69 (74.6)	3.84 (60.5)	4.00 (60.5)		
Hex3	Gal I	*β*-D-Gal*p*-(1	N/R	4.47 (103.1)	3.55 (74.8)	3.68 (73.3)	3.94 (68.7)	3.75 (75.5)	3.74 (62.8)	N/D		
Hex4	Gal I	4)-*β*-D-Gal*p*-(1	7.5	4.53 (103.3)	3.59 (70.7)	3.76 (72.3)	4.06 (77.6)	3.80* (75.5)	3.86 (60.5)	3.92 (60.5)		
Hex4	Gal II	*α*-D-Gal*p*-(1	3.7	4.97 (102.1)	3.84 (68.7)	3.90 (69.3)	4.04 (69.2)	4.35* (71.0)	3.72 (60.8)			
all	PEtn			4.15 (62.1)	3.24 (40.8)							
all	PCho			4.38 (59.6)	3.69 (66.2)							

Hex3 bears Gal*β*(1,4)Glc*β*(1,2) at Hep III, while Hex4 bears Gal*α*(1,4)Gal*β*(1,4)Glc*β*(1,2) at this position. $^{3}J_{H,H}$ values for anomeric ^{1}H resonances (H1) are given. Except where indicated, chemical shifts are the same (± 0.02 ppm) for the conserved inner core residues, shared by both glycoforms. N/R: not resolved, N/D: not defined. *Tentatively assigned from ROESY spectra. †Calculated from TOCSY spectra.

Hex4 glycoform. The presence of the globotriose moiety at Hep III was confirmed by the ROE connectivities between H1 of Gal II and H4/H6 of Gal I (Fig. 4.15 and Table 4.5).

Table 4.5 Proton ROE data for NTHi375Δ*ompP5*-derived OSs

Anomeric proton	Observed proton ROE	
	Intra-residue	**Inter-residue**
5.09-5.14 (Hep I)	4.05 (H2)	–
5.60-5.68 (Hep II)	4.32 (H2)	3.70 (Hep I H3); 4.19; 4.21 (Hep III H2)
5.10-5.13 (Hep III)	4.19; 4.21 (H2)	5.60-5.68 (Hep II H1); 4.32 (Hep II H2); 4.68 (Glc II H1)
4.52 (Glc I)	3.48 (H3) [a] and 3.60 (H5)[a]	4.25 (Hep I H4); 4.14 ($H6_A$)
4.68 (Glc II)	3.69 (H3/H5)[b]	5.10-5.13 (Hep III H1); 4.23 (H2)
4.53 (Gal I)	3.76 (H3) and 3.80 (H5)	3.69 (Glc II H4)[b]
4.97 (Gal II)	3.84 (H2) and 3.90 (H3)	4.06 (Gal I H4); 3.86 ($H6_A$); 3.92 ($H6_B$)

[a] Similar intra-residue distance of ~2.5 Å from the anomeric proton;
[b] ^{1}H-signals corresponding to H3, H4 and H5 of Glc II were overlapping (~3.69 ppm), resulting in a broadened ROE cross-peak (also see Fig. 4.15).

Based on the relative ^{1}H-^{13}C HSQC intensities of the anomeric signals of terminal Gal*p* in the Hex3 and Hex4 glycoforms (Gal I and II respectively), Hex3 was roughly estimated to represent a population of ~22% of total OSs. NTHi375 LOS was previously described to be predominantly constituted by Hex4 glycoforms, together with mono- and di-sialylated Hex3 species representing 9% and 5% of the LOS, as determined by electrospray-ionisation mass spectrometry (Fox et al, 2006). No OS-bound Neu5Ac was detected in our NMR analysis. At first place, the small difference in population reported for total sialylated species compared to that estimated here for Hex3 could be attributed to the different detection methods used in each case. Besides, it seems highly likely that the conditions used for cleavage of the Kdo-lipid A linkage could also result in hydrolysis of Neu5Ac residues, yielding the Hex3 glycoform here observed. Indeed, the ^{1}H-NMR spectra of Neu5Ac before and after treatment at our acid hydrolysis conditions (section 3.8.1) differed substantially, the appearance of new peaks pointing to further hydrolysis of the Neu5Ac moiety besides the cleavage of the α(2,3)-glycosidic bond (Annex II). Thus, the Hex3 OS would correspond to desialylated species.

A similar set of NMR analyses was carried out for the Rd OSs (Annex II). The results obtained were comparable to those previously reported for this strain (Risberg *et al.*, 1999) confirming the presence of three main glycoforms bearing terminal *β*Gal (Hex3), *α*Gal (Hex4), or *β*GalNAc (Hex5) at the Hep III branch (see Fig. 4.11).

4.1.7.2 The LOS epitopes recognized by VAA and RCA are lectin- and OS-specific

Following OSs assignment, we proceeded to the study of their recognition by the lectins. As no significant binding was detected for RCA in the NTHi375Δ*ompP5* LOS microarrays (Fig. 4.9), the STD analysis was carried out only for VAA in this case. Aromatic irradiation (7 ppm) provided sufficient protein saturation (82%), while aliphatic radiation (-0.5 ppm) was less efficient, with protein saturation barely reaching 70%. Control experiments of the NTHi375Δ*ompP5*-derived OSs were performed in the absence of VAA in order to ensure that its protons did not suffer non-specific saturation upon irradiation.

Under the experimental conditions followed (1:30 VAA:ligand ratio), the highest absolute STD signal observed, corresponding to Gal II H4, was 1.8% compared to the

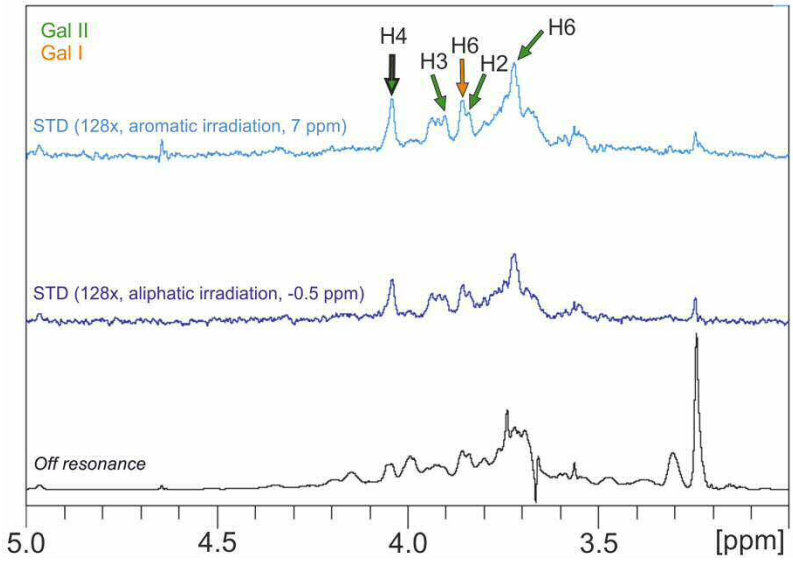


Fig. 4.16 STD spectra of the VAA- NTHi375Δ*ompP5* OS complex.
STD NMR spectrum obtained upon irradiation at 7 and -0.5 ppm. A lower saturation of the receptor during aliphatic irradiation yielded a poorer saturation transfer to the ligand, compared to aromatic irradiation. Both spectra indicate the ligand protons that interact stronger with VAA, following an identical profile. The resonance corresponding to H4 of the terminal *α*-Gal-moiety gave the most intense STD signal. This and other STD signals are highlighted in the top spectrum. The bottom spectrum corresponds to the *off-resonance* spectrum of a 1:30 VAA:NTHi375Δ*ompP5* OS mixture irradiated at 100 ppm after protein signal subtraction.

Table 4.6 STD intensities of Gal II and Gal I in the Hex4 glycoform of NTHi375Δ*ompP5* OS

	Gal II [*α*-D-Gal*p*-(1→]			**Gal I** [→4)-*β*-D-Gal*p*-(1→]		
	^{1}H	**STD**	**Normalised[a] STD**	**^{1}H**	**STD**	**Normalised[a] STD**
H1	4.97	N/D[b]	N/D[b]	4.53	ND[b]	ND[b]
H2	3.84	0.95%	56%	3.59	1%	56%
H3	3.90	1.5%	83%	3.76	0.9%	50%
H4	4.04	1.8%	100%	4.06	0.9%	50%
H5	4.35	1.75%	95%	3.80	0.9%	50%
$H6_{A-B}$	3.72	1.2%	67%	3.86; 3.92	1.2%	67%

[a] Normalised values were calculated taking the signal of Gal II H4 as 100%; [b] shift close to residual HDO signal, unreliable STD quantitation.

off-resonance spectrum after protein background signal subtraction (Fig. 4.16 and Tables 4.4 and 4.6). This STD signal was normalised to 100%. For each proton whose signal intensity could be individually measured, its STD value was obtained and normalised to the most intense signal (Table 4.6). Unfortunately, the use of a pulse sequence with water suppression (WATERGATE) affected the resonances around 4.7 ppm, so no STD values could be reported for the anomeric protons of the two Gal moieties potentially involved in binding. In addition, for Gal II the $H6_A$ and $H6_B$ signals overlapped (Table 4.4), so only a mean STD value could be determined in this case (Table 4.6). The lower protein saturation achieved during irradiation at -0.5 ppm resulted in lower saturation transfer for Gal II, the H4 signal being affected only by 1.3%. However, the spectral profile was identical to that observed upon aromatic irradiation (Fig. 4.16), confirming the proximity of the defined protons to the protein.

Major contacts engaged the *α*-terminal Gal residue of the major Hex4 glycoform, while saturation transfer to protons of the inner *β*-Gal was also observed. Since Gal II H4, H5 and H3 received maximal saturation (Fig. 4.17) it can be deduced that these are the

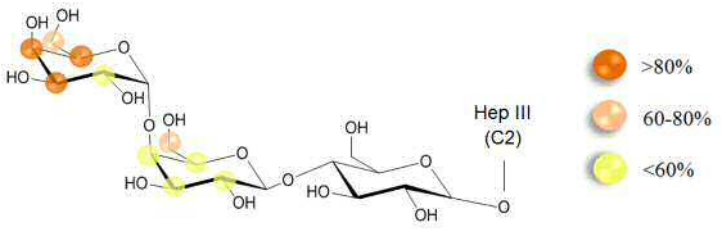


Fig. 4.17 STD-based epitope mapping for VAA binding to the globotriose moiety of NTHi375Δ*ompP5* OS. Percentages upon normalisation to the most intense signal, ascribed to H4 of the terminal *α*-Gal (Gal II).

protons closer to the protein. The crystallographic structures of VAA-Gal (PDB: 1PUM, 1OQL) and VAA-Lac (PDB: 1PUU) complexes show that in the binding site operative at the VAA concentrations used for the STD-NMR experiments (Jiménez *et al.*, 2005; Jiménez *et al.*, 2006) the phenolic ring of a Tyr residue (Tyr249) participates in CH-π stacking with the Gal*p* ring, a frequent feature in lectin binding sites (Asensio *et al.*, 2013), engaging protons H3, H4 and H5 (Fig. 18) (Fernández-Alonso *et al.*, 2005). Thus, considering the proximity of these protons to Tyr249, which was probably directly irradiated, high saturation transfer to the three of them was expected. In addition, the hydroxyl groups at positions 3 and 4, the latter defining the Gal stereochemistry, are involved in strong hydrogen bonds with the protein (Fig. 4.18).

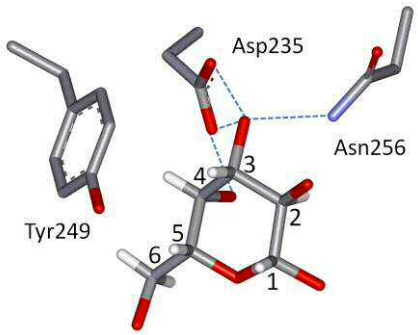


Fig. 4.18 Stacking interaction and key hydrogen bonds observed in the crystal structure of the VAA-Gal complex. Positions in the Gal moiety are numbered and the key hydrogen bonds established by the hydroxyl groups at positions 3 and 4 are represented by dotted lines.

On the other hand, although HO-6 is not involved in the binding (Jiménez *et al* 2008), the H6 protons can be also oriented towards the Tyr side chain, so that saturation transfer to these protons appears possible. Worth to mention, STD values for *α*-methyl-galactoside H6 in presence of VAA were the highest upon receptor saturation for 0.5 s, yet around 75% for longer saturation intervals (El Biari, 2014). Saturation of approx. 2 s was pursued throughout the STD experiments here presented. Finally, a STD signal was also observed for Gal II H2, with intensity comparable to that observed for the Gal I protons. The HO group at position 2 has been proposed to be involved at neutral pH in a water-mediated interaction with the protein (Jiménez *et al.*, 2008), indicating that this position is still accommodated within the contact area. The STD signals observed for Gal I protons hint at possible medium-distance contacts of this unit with the protein. Of note, a previous study of the binding of the Gal*α*(1,4)Gal disaccharide to VAA also revealed STD signals for the reducing Gal residue (Miller *et al.*, 2011), supporting the results here reported.

Worth mentioning, a watchful look to the STD spectrum of VAA–NTHi375Δ*ompP5* OS (Fig. 4.16) suggests that more protons may suffer saturation transfer and be involved in binding, including those belonging to the terminal *β*-Gal of the Hex3 glycoform, in agreement with the lactose epitope also serving as ligand for VAA (Jiménez *et al* 2008). Yet, a quantitative analysis and normalisation of these signals could be erratic and was not therefore attempted, especially considering that the Hex3 glycoform only appears as the result of hydrolysis of the sialylated LOS species and is not naturally present in the NTHi375Δ*ompP5* LOS. Indeed, the absence of a terminal Lac epitope explains the lack of binding of RCA to this LOS in the microarray binding assays, since this lectin exhibits a strong preference for *β*-linked over *α*-linked Gal (Podder *et al.*, 1974; Baenziger and Fiete, 1979; Wu *et al.*, 1993).

In contrast, strong binding of RCA to the Rd LOS was observed (Fig. 4.12). As already mentioned, the major glycoform of the Rd LOS presents terminal *β*Gal(1,4)Glc (Fig. 4.11), which could serve as docking point for RCA. To confirm this postulation, the binding of RCA to the Rd-derived OSs was examined (Fig. 4.19). STD for *β*-Gal (Gal I) H4 corresponded to 2% of the initial *off-resonance* signal. Other affected protons included H2 (normalized STD value of 98%), H3 (85%) and $H6_{A-B}$ (60%) of this residue, followed by the overlapping H3/H4 and H5 of the proximal Glc II moiety (mean value of 40%), Table 4.7 and Fig. 4.19. All these signals were fully consistent with the differential contribution of the HO groups at these positions to the binding of lactose to RCA, as determined by chemical mapping analysis (Rivera-Sagredo *et al.*, 1992; Solís *et al.*, 1993). The particularly high contribution of Gal I H2, could be attributed to the establishment of a direct hydrogen bond with lectin residues, as a

Table 4.7 STD intensities of Gal I and Glc II in the Hex3 glycoform of the Rd OS.

	Gal I [*β*-D-Gal*p*-(1→]			**Glc II** [→4)-*β*-D-Glc*p*-(1→]		
	^{1}H	**STD**	**Normalised[a] STD**	**^{1}H**	**STD**	**Normalised[a] STD**
H1	4.47	N/D[b]	N/D[b]	4.67	N/D[b]	ND[b]
H2	3.55	1.95%	98%	3.38	0.55%	28%
H3	3.68	1.7%	85%			
H4	3.94	2%	100%	3.69[c]	0.8%	40%
H5	3.74	0.6%	30%			
$H6_{A-B}$	3.78; N/D	1.2%	60%	3.85; 3.99	0.7%;0.5%	35%; 25%

[a] Normalised values were calculated taking the signal of Gal I H4 as 100%; [b] shift close to residual HDO signal, unreliable STD quantitation. [c] signal overlapping, the mean STD value is given.

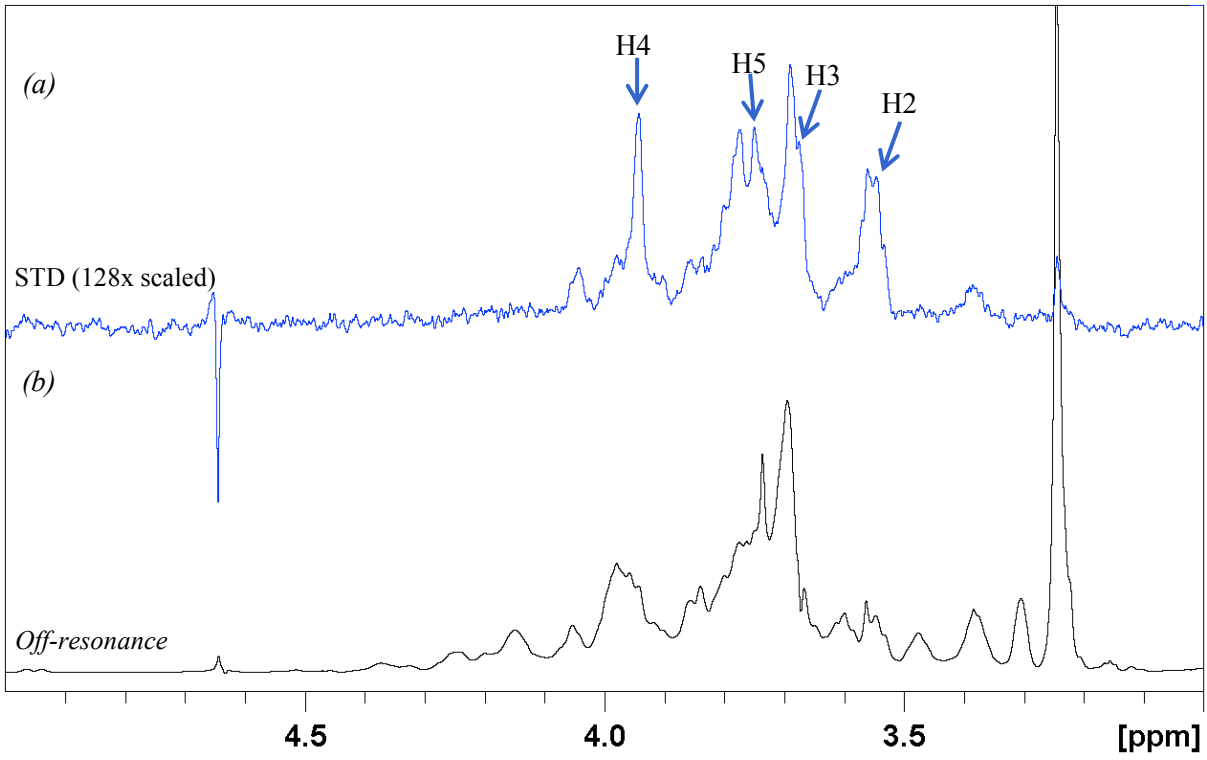


Fig. 4.19 STD and reference spectra of the RCA-OS NTHi Rd complex.

(a) STD NMR spectrum revealing the directly interacting OS protons. The resonance corresponding to the H4 of the terminal *β*-Gal I, gave the most intense STD signal. Other significant STDs are denoted by arrows and reported in Table 4.7. *(b)* Section of an *off-resonance* spectrum of a mixture of RCA and OS from NTHi375Δ*ompP5* in a 1:30 ratio, after protein signal subtraction. RCA was irradiated at 100 ppm.

donor, as previously suggested (Bhattacharyya and Brewer, 1988). No meaningful signals for other sugar protons were observed, also in agreement with the reported RCA binding specificity (Podder *et al.*, 1974; Baenziger and Fiete, 1979; Wu *et al.*, 1993), including its inability to accommodate GalNAc residues at the binding site. Thus, recognition of the Hex3 glycoform accounts for the binding of RCA to the Rd LOS.

A similar analysis of VAA binding yielded similar STD signals (Fig. 4.20). Again, the terminal *β*-Gal H4 proton (3.94 ppm) gave the strongest signal with an absolute STD value of 2%. Saturation transfer was also effective to the H3, H5, and H6 of Gal I. Although the poor signal-to-noise ratio precluded a detailed quantitative analysis, the results evidenced that VAA is also able to recognize the Hex3 glycoform, here representing ~60% of total OSs (Risberg *et al.*, 1999). The absence of meaningful signals for *α*Gal was intriguing, as the Hex4 glycoform, serving as ligand for VAA in the NTHi375Δ*ompP5* LOS, accounts for nearly 20% of the Rd OSs. Interestingly, colony immunoblotting had previously revealed weak or even undetectable binding of the Gal*α*(1,4)Gal-specific monoclonal antibody 4C4 (High *et al.*, 1993) to this strain

(Virji *et al.*, 1990), what could indicate a very different presentation of this epitope in the Rd LOS that is neither appropriate for VAA binding.

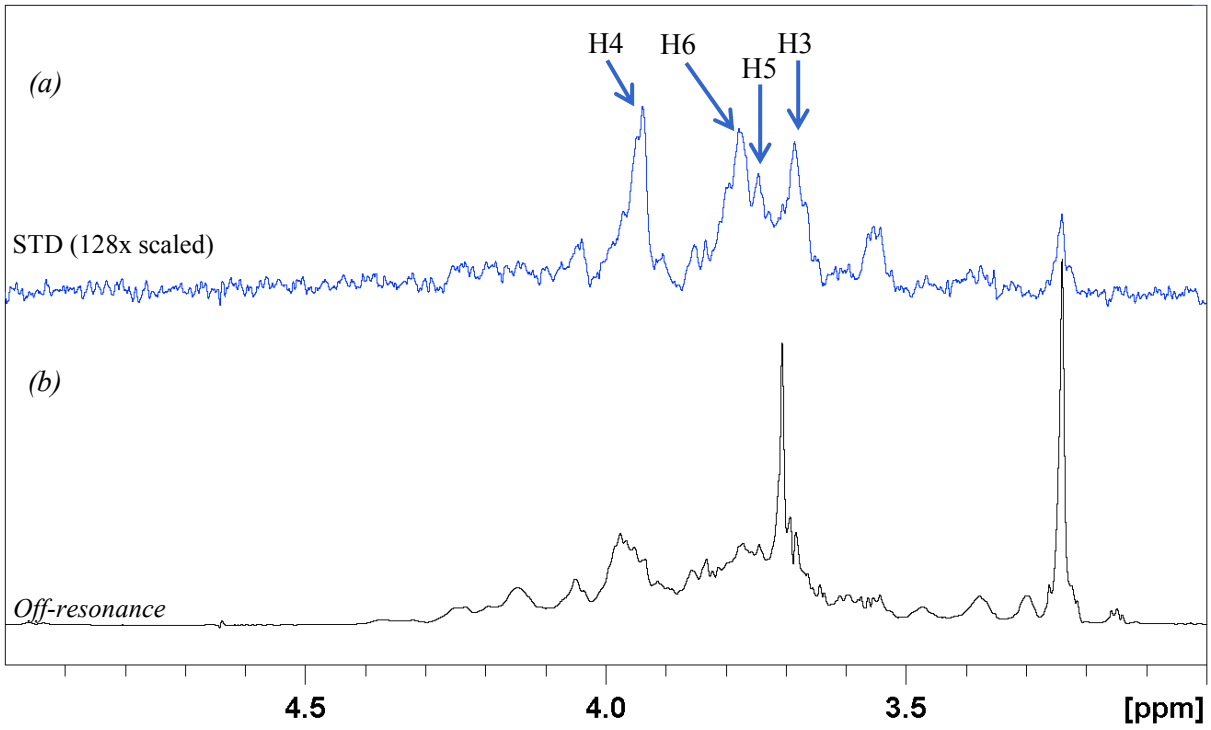


Fig. 4.20 STD and reference spectra of the VAA-OS NTHi Rd complex.
(a) STD NMR spectrum revealing the directly interacting OS protons. The resonance corresponding to the H4 of the terminal *β*-Gal I, gave the most intense STD signal. Other significant STDs are denoted by arrows. *(b)* Section of an *off-resonance* spectrum of a mixture of VAA and OS from NTHi375Δ*ompP5* in a 1:30 ratio, after protein signal subtraction. VAA was irradiated at 7 ppm.

4.1.8 Profiling of accessible glycans on NTHi clinical isolates and recognition by endogenous lectins

A hallmark of NTHi is its inter-strain genomic and phenotypic heterogeneity (Martí-Lliteras *et al.*, 2011; De Chiara *et al.*, 2014). The LOS outer core is highly variable due to the presence/absence and phase variation of genes responsible for initiating extensions from the three inner core Hep residues (Hood *et al.*, 2004), leading to variable outcomes as on pathogen colonisation, persistence, or acute infection (Schweda *et al.*, 2007). Prompted by the demonstrated utility of NTHi microarrays for profiling accessible glycans on the bacterial surface, we next compared the lectin-binding patterns of six different NTHi clinical isolates. We selected four genotypically distinct clinical isolates of two frequent pathological origins, otitis media (strains 375 and 86-028NP) and COPD (strains 398 and 1566). These strains had been previously genome sequenced (Table 3.2) and found to present heterogeneous genomic contents. In

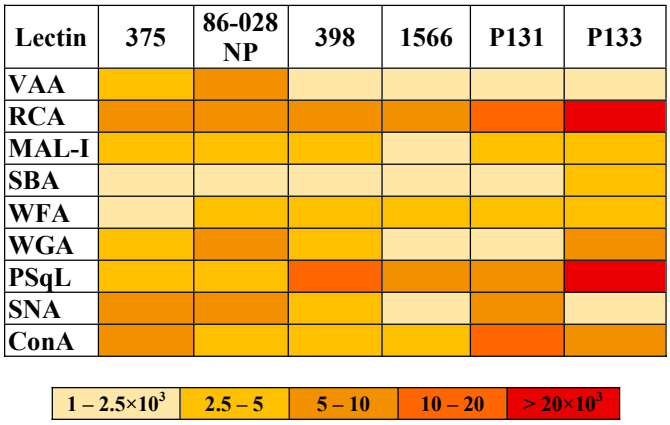


Fig. 4.21 Clinical isolates were printed in the arrays, and the binding of selected lectins was examined. The colour code reflects the fluorescence intensity of AF647-streptavidin signals (in rfu) used for detecting the biotinylated lectin binding to OD_{600} = 1 of printed bacteria in the microarrays. Detailed data are given in Table 4.8.

addition, we selected two NTHi isolates from paediatric healthy carriers (strains 131 and 133), which exhibited different pulse field gel electrophoresis (PFGE) profiles (personal communication of Dr. Junkal Garmendia). As illustrated in Fig. 4.21 and Table 4.8, the "lectin fingerprint" obtained for each isolate was different.

A clear-cut dissection and assignment of the glycoepitopes recognised by the lectin panel is not feasible, particularly considering that detailed information on the LOS structure for all NTHi strains is not available. Still, the lectin-binding fingerprints of strains 131 and 133 deserve to be highlighted as they yielded the strongest signals for key lectins.

PSqL gave strong signals for both isolates, binding to 133 giving the most intense signal among all strain–lectin pairs tested (Table 4.8). A high content of sialic acid-bearing ligands could be tentatively connected with a role in allowing the microbe to masquerade as "self" and thereby to either elude or subvert host immune responses (Chang and Nizet, 2014), at the same time contributing to biofilm formation and persistence (Swords *et al.*, 2004). Binding signals observed for RCA also revealed a

Table 4.8 Binding of selected reference lectins to NTHi clinical isolates[a]

	Otitis media		COPD		Healthy carriers	
Lectin	375	86-028NP	398	1566	P131	P133
VAA	4250 ± 150	5050 ± 60	1730 ± 20	1850 ± 150	2100 ± 100	2300 ± 300
RCA	5100 ± 700	5500 ± 400	8700 ± 700	6800 ± 700	10400 ± 400	25500 ± 900
MAL-I	2800 ± 150	2870 ± 80	2900 ± 200	1290 ± 75	2900 ± 150	2950 ± 200
SBA	2000 ± 100	2400 ± 400	2250 ± 450	1700 ± 550	1900 ± 50	2500 ± 900
WFA	2450 ± 100	3100 ± 150	3850 ± 300	2700 ± 450	3200 ± 500	4500 ± 750
WGA	3300 ± 250	5450 ± 200	4000 ± 300	1100 ± 150	1700 ± 150	5600 ± 300
PSqL	3700 ± 200	2900 ± 200	13000 ± 600	5600 ± 300	9000 ± 800	28000 ± 4800
SNA	5050 ± 200	6050 ± 300	3850 ± 200	1870 ± 80	6150 ± 450	2200 ± 250
ConA	5090 ± 90	4300 ± 150	3000 ± 150	3900 ± 200	11100 ± 200	5200 ± 150

[a]Data shown corresponds to mean values of the fluorescence intensity of AF647-streptavidin signals (in arbitrary units) at the highest concentration of bacteria tested (OD_{600} = 1), calculated from the intensities obtained for triplicate samples of two different experiments. Standard deviations of the mean are given.

high availability of ligands for this lectin. In contrast, binding of VAA was noticeably smaller, again evidencing that RCA and VAA recognise different Gal-containing ligands.

Of note, the two isolates from otitis media presented a greater abundance of VAA ligands compared to the rest of isolates. LOS capping by *α*-Gal has been previously found to contribute to evasion of acquired immunity, also being advantageous in inflammatory states (Weiser and Pan, 1998). However, in the case of COPD isolates only weak binding of VAA was detected, suggesting a low availability of the digalactose Gal*α*(1,4)Gal epitope. These results were in line with blotting assays of bacterial lysates showing no reactivity of strains 398 and 1566 towards the Gal*α*(1,4)*β*Gal-specific 4C4 monoclonal antibody, although these strains carry the phase-variable gene *lic2A* in frame (Martí-Lliteras *et al.*, 2011). Noteworthy, NTHi1566 in particular also gave the smallest binding signals for other lectins, pointing to the unavailability of other epitopes.

At any rate, the microarray lectin-binding assays evidenced the availability on the bacterial surface of Man/Glc, Gal and Neu5Ac residues that might serve as docking points for lectins of the innate immune system. To confirm this hypothesis, we took advantage of our microarray platform to analyse the binding to the isolates of SP-D, hGal-8, and Siglec-14, as representative lectins with the required nominal binding specificities (see Table 1.2).

SP-D is a calcium-dependent lectin of the collectin family involved in surfactant homeostasis and pulmonary immunity (Waters *et al.*, 2009; Seaton *et al.*, 2010). Collectins recognize pathogen-associated molecular patterns of a variety of microorganisms, most of them exhibiting mannose-like binding specificity (Gupta and Surolia, 2007). Galectins have also been shown to participate in defence against microbial pathogens, through regulation of innate and adaptive immunity (Rabinovich *et al.*, 2012). In addition, they interact directly with pathogens leading to either enhancement or blockage of microbial infection, or to microbicide activity (Baum *et al.*, 2014; Chen *et al.*, 2014). In particular, a direct microbicidal effect of hGal-8 on NTHi2019 was demonstrated (Stowell *et al.*, 2014). A variety of interactions with bacterial, viral and protozoan pathogens have been described for Siglecs too, which also have important innate and adaptive immune functions (Chang and Nizet, 2014;

Macauley *et al.*, 2014). Binding of NTHi2019 to Siglec-14, expressed on neutrophils and monocytes, was found to enhance inflammation, triggering COPD exacerbation (Angata *et al.*, 2013). Of note, compared to NTHi375, the NTHi2019 LOS has, as distinctive structural features, the presence of Galβ(1,4)Glcβ(1,4) linked to the proximal Hep I and of Galβ(1,4)Glcβ(1,4)Glcα(1,3), and truncated versions thereof, linked to the middle Hep II (Phillips *et al.*, 1992). This poses the question of the extrapolability of Siglec-14 and, especially, hGal-8 binding to NTHi375 as well as to other NTHi strains.

Analysis of the binding to the array-printed NTHi clinical isolates revealed a dose-dependent and strain-specific recognition by the three lectins (Fig. 4.22), providing the first experimental evidence for direct binding of SP-D to NTHi, and also demonstrating the binding of hGal-8 and Siglec-14 to NTHi strains other than NTHi2019 (Kalograiaki *et al.*, 2016). Although the different binding-detection protocols used for the three lectins (please see Fig. 3.12) make inappropriate an inter-lectin comparison of binding intensities, the strain-specific binding patterns observed for each lectin can be confidently compared. Once again, the smallest binding signals were observed for

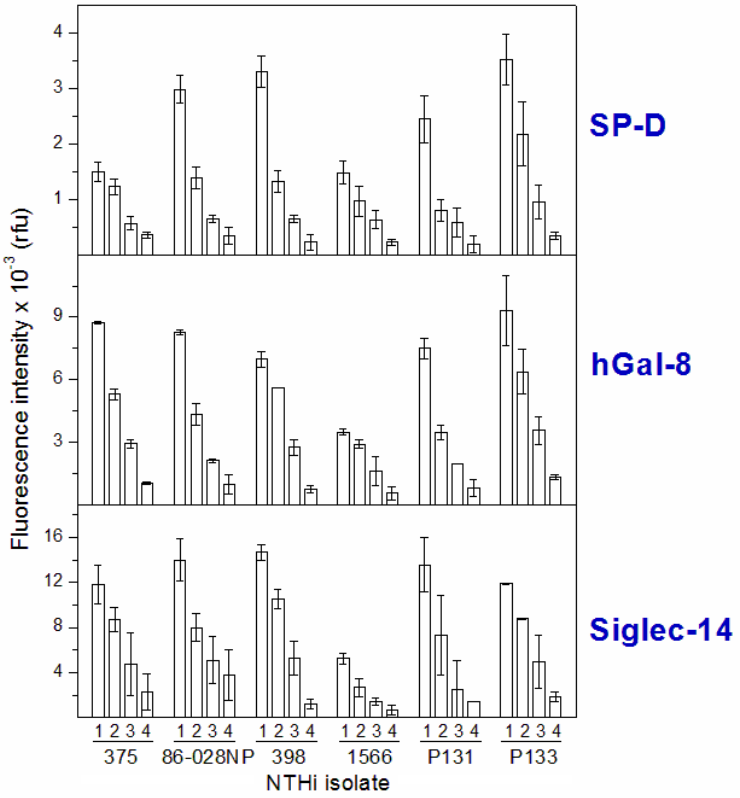


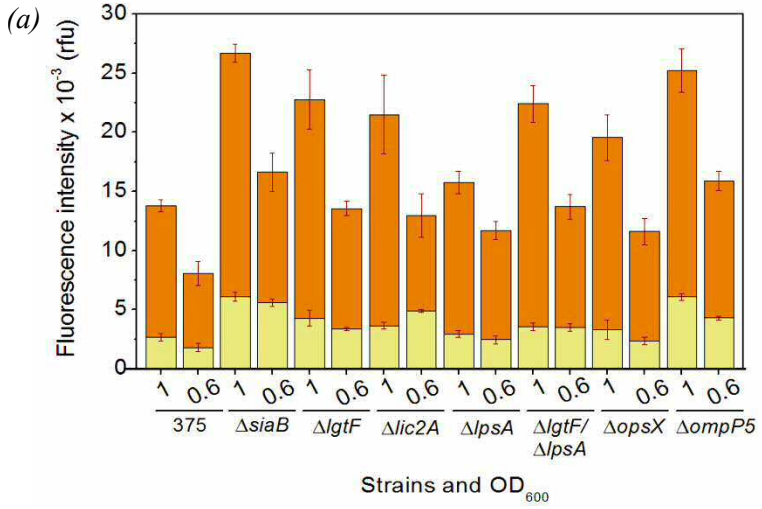
Fig. 4.22 Binding of innate immune lectins to NTHi clinical isolates. Bacteria strains were printed as triplicates at four dilutions (1-4, corresponding to OD_{600} of 1, 0.6, 0.3 and 0.1, respectively). Slides were overlaid with either SP-D, hGal-8 or Siglec-14, and binding was detected by incubation with AF647-streptavidin as final step. Dotted lines represent the observed binding trends, distinct for each lectin.

NTHi1566. For the other isolates tested, lectin-specific trends were observed (Fig. 4.21) indicating that the availability of the respective epitopes varies from strain to strain.

Interestingly, only in the case of SP-D the binding to strain 1566 was comparable to that detected for NTHi375. SP-D exhibits a broad binding-specificity with maltose (Glcα(1,4)Glc), glucosides and mannosides with a preference for α-anomers (Persson *et al.*, 1990) , as well as L,D-manno-Hep (Wang *et al.*, 2008; Clark *et al.*, 2016) being recognised ligands. As illustrated in Fig. 4.1, the NTHi LOS contains an Hep trisaccharide inner core, which is conserved in several bacterial species (Hood *et al.*, 2004) and could serve as docking point for SP-D. Therefore, prompted by the observed strain-selective binding of SP-D to all NTHi clinical isolates, we explored the recognition by this lectin of the library of NTHi375 mutants displaying sequentially truncated LOS, with Hep residues cloaked or exposed to different degrees.

4.1.9 Recognition of NTHi375 mutants and isolated LOSs by SP-D

SP-D binding to microarray-printed NTHi375Δ*siaB*, Δ*lgtF*, Δ*lgtF*/Δ*lpsA* and Δ*ompP5* strains was significantly enhanced compared to the wild type strain, while deletion of the Hep III extension led to a moderate binding increase (Fig. 4.23). Hence, increasedexposure of the LOS inner core gave rise to a boost in SP-D binding (Fig.



(b)

Mutant	Relative SP-D binding
Δ*siaB*	200 ± 20
Δ*lgtF*	170 ± 11
Δ*lic2A*	160 ± 12
Δ*lpsA*	129 ± 6
Δ*lgtF/lpsA*	180 ± 11
Δ*opsX*	150 ± 10
Δ*ompP5*	210 ± 10

Fig. 4.23 Binding of SP-D to the panel of NTHi375 mutants.
(*a*) Representative binding of 6 μg/ml SP-D to immobilised bacteria (triplicates of OD_{600} values 1 and 0.6) in the absence (orange) or presence of 0.2 M maltose (yellow). Data shown correspond to the mean of three data points obtained in the same microarray slide, upon normalisation for bacterial staining with SYTO® 13. Error bars indicate the standard deviation of the mean. (*b*) Relative binding strength calculated as percentage taking binding to wild type NTHi375 as 100%, upon normalisation for staining with SYTO® 13. The colour code (light to dark) denotes the relative binding increase. Data correspond to at least two independent experiments. Standard error values are given.

4.23, *b*). Indeed, it has been reported that SP-D prefers rough LPS over smooth LPS as ligand (Kuan *et al.*, 1992; Ofek *et al.*, 2001), probably because of a higher accessibility to the inner core Hep (Wang *et al.*, 2008).

The specificity of the binding was assessed using different competitors (Table 4.9). First, the binding to all strains was partially inhibited (30-56%) in the presence of 0.05 M L,D-manno-Hep and to a higher extent (up to 70-77%) in the presence of 0.2 M maltose (Fig. 4.23, *a*). In contrast, Kdo was not a competitor when used at concentrations of up to 50 mM (not shown), in agreement with previous observations (Wang *et al.*, 2008; Reinhardt *et al.*, 2016). In addition, since Ca^{2+} is essential for SP-D's lectin-binding activity (Shrive *et al.*, 2003; Kishore *et al.*, 2006), the effect of EDTA was also tested, inhibition reaching 80-89% of the binding in the presence of Ca^{2+}. Altogether, the results evidenced that SP-D binding to the NTHi strains was carbohydrate-mediated.

Table 4.9 Inhibition of SP-D binding to NTHi375 wild type and mutant strains by different competitive haptens[a]

Strain	Inhibitor		
	Maltose	Manno-Hep	EDTA
NTHi375	76 ± 5	32 ± 5	88 ± 5
Δ*siaB*	72 ± 4	41 ± 3	81 ± 4
Δ*lgtF*	69 ± 2	39 ± 9	80 ± 3
Δ*lic2A*	75 ± 5	30 ± 3	85 ± 4
Δ*lpsA*	71 ± 3	51 ± 3	81 ± 5
Δ*lgtF/lpsA*	77 ± 4	56 ± 4	89 ± 5
Δ*opsX*	76 ± 3	35 ± 2	84 ± 3
Δ*ompP5*	70 ± 3	41 ± 3	87 ± 3

[a]SP-D binding was tested at 6 μg/ml. Hapten concentrations were 0.2 M maltose, 0.05 M manno-Hep, and 0.01 M EDTA. Data correspond to the mean of at least two competition experiments. Standard error values are given.

Binding to Δ*opsX* (exposing an unsubstituted Kdo moiety) was slightly lower compared to Δ*lgtF*. This finding was in good agreement with a weaker SP-D binding to a deep rough Re-LPS *E. coli* mutant bearing a terminal α(2,4)-linked di-Kdo moiety, compared to its Rd form substituted in position 5 with 4-phosphoryl-L,D-manno-Hep, akin to NTHiΔ*lgtF* (Wang *et al.*, 2008), revealing a preference for terminal Hep over Kdo. More important, the efficient binding to NTHi375Δ*siaB* and Δ*ompP5* implies that, if the

inner core of the NTHi375 LOS is a docking point for SP-D on the bacterial surface, this lectin should be able to recognise internal manno-Hep, as previously suggested (Wang *et al.*, 2008; Reinhardt *et al.*, 2016). To further clarify this issue, a direct comparison of SP-D binding to the LOSs isolated from the Δ*ompP5*, Δ*lpsA* and Δ*lgtF* mutants was carried out.

To this aim, the LOS of the Δ*lgtF* and Δ*lpsA* mutants was extracted and quantified as described above (section 4.1.5.1). A higher extraction yield was again observed compared to wild type NTHi375 LOS (Fig. 4.24), what in this case could result from a differential partition index in the extraction media due to structural differences. Interestingly, a second weak band was detected for the Δ*lgtF*-derived LOS (highlighted by an arrow in Fig. 4.24), which apparently corresponds to a shorter isoform with mobility close to that of the Δ*lpsA* LOS. Of note, this mutant strain was found to be reactive to the anti-PCho antibody TEPC-15, suggesting that in the absence of Glc I at Hep I, PCho could be added to Glc II, linked to Hep III (Morey *et al.*, 2013). Thus, a tentative explanation for the appearance of this second band could be that PCho conjugation to Glc II in a small population of the Δ*lgtF* LOS could be a source of spatial hindering, preventing further branch extension.

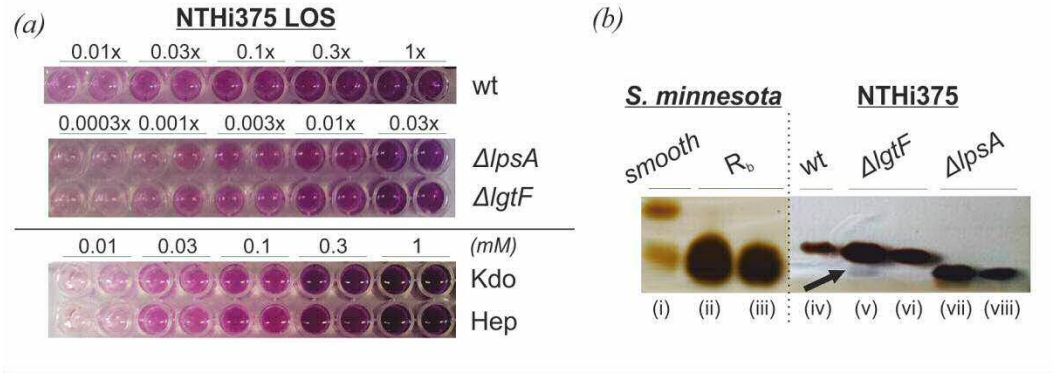


Fig. 4.24 Quantitation of the LOS isolated from NTHi375 Δ*lgtF* and Δ*lpsA* mutant strains. *(a)* Colorimetric quantitation by the Purpald assay. LOS dilutions from the stock solution are indicated (in duplicates). *(b)* Electrophoretic profiles of the isolated LOSs, sample loaded: (i) 1 µg, (only shown the two lower M_w bands); (ii) 1 µg, single band; (iii) 0.7 µg; (iv) 3 µl stock; (v) 1 µl (dil 1:200); (vi) 0.6 µl (dil 1:200); (vii) 1 µl (dil 1:200); (viii) 0.6 µl (dil 1:200). The arrow indicates the presence of a second weak band in the LOS of Δ*lgtF* (v).

Quantified Δ*lgtF* and Δ*lpsA* LOSs were incorporated in the microarrays at equal concentrations, along with Δ*ompP5* and *S. minnesota* Rb LOS as references. The arrays were validated by confirming binding of the polyclonal anti-*Salmonella*-lipid A antibody, as described in section 4.1.5.1. A slight tendency to an increase in the binding

intensity with LOS truncation was observed (Fig. 4.25, *a*), what could be attributed to a higher lipid A accessibility. Of note, no binding of mouse polyclonal anti-*Salmonella typhimurium* antibody to the printed NTHi LOSs was detected, supporting the specificity of the recognition by the anti-lipid A antibody (not shown). The binding of SP-D to Δ*ompP5,* Δ*lgtF* and Δ*lpsA* LOSs printed in the arrays was then examined in the absence and presence of 0.2 M maltose or 0.05 M L,D-manno-Hep.

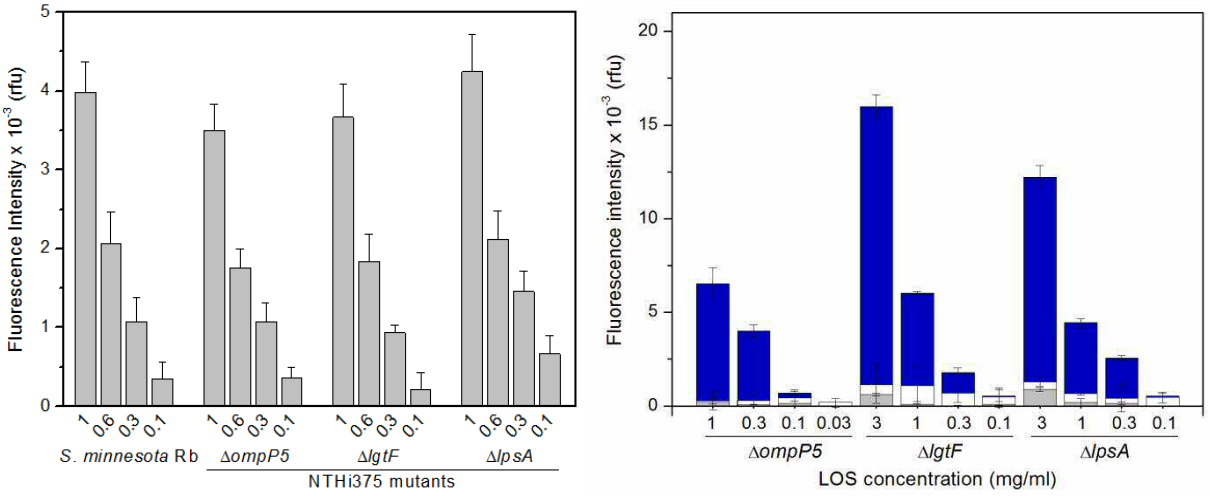


Fig. 4.25 Binding of anti-lipid A antibody and SP-D to LOS microarrays. The LOSs were immobilised at decreasing concentrations as triplicates, and the binding of anti-lipid A (left panel), and SP-D was examined (right panel). For SP-D, the binding was tested in the absence (blue) or presence of 0.2 M maltose (grey) or 0.05 M L,D-manno-Hep (white). Error bars stand for standard deviation of the mean.

Efficient binding to the LOS isolated from Δ*ompP5* was detected (Fig. 4.25, *b*). Moreover, the binding was comparable to that to Δ*lgtF* and Δ*lpsA* LOS at the same printed concentrations, evidencing that exposure of terminal Hep I or Hep III is not necessary for recognition. In all cases, the binding was inhibited almost completely in the presence of competitive haptens (Fig. 4.25, *b*), proving that it was entirely carbohydrate-mediated, without the involvement of hydrophobic protein-lipid interactions. Thus, incomplete inhibition by the tested haptens of SP-D binding to the whole bacteria could attest for the existence of alternative high-affinity docking points for the lectin, or alternatively reflect the establishment of secondary carbohydrate–independent interactions.

4.1.10 Recognition of NTHi375 mutants and isolated LOSs by galectins

Galectins constitute an attractive lectin family involved in multiple processes related to health and disease. Their CRDs are highly conserved, with a common sequence signature that determines their specificity for *β*-galactosides (Annex IV). However,

small variations in the sequence and architecture of the carbohydrate-binding site lead to distinct fine oligosaccharide-binding specificity, particularly towards Gal*β*(1,4)GlcNAc- Gal*β*(1,3)GlcNAc- or Gal*β*(1,3)GalNAc-containing oligosaccharides and sialylated species thereof. What is more, their structural organisation defines three groups, i.e. proto, chimera and tandem-repeat type (Fig. 4.26, *a*), with distinct capacities to bind ligand clusters (*cis*-crosslinking) and/or act as bridges (*trans*-crosslinking) (Fig. 4.26, *b*). Galectins are present in all tissues, but certain mucosal epithelia (*e.g.* Gal-4 in intestine or Gal-8 in larynx), show a particularly high expression, what appears to be associated to crucial physiological roles, including pathogen recognition and clearance. The linker peptide connecting the two different CRDs of tandem-repeat galectins (Fig. 4.26, *a*), which is composed of a minimum of 20 amino acids, is susceptible to proteolysis (Oda *et al.*, 1993), and has been proposed to govern the CRDs' proximity, relative orientation and crosslinking capacity (Earl *et al.*, 2011).

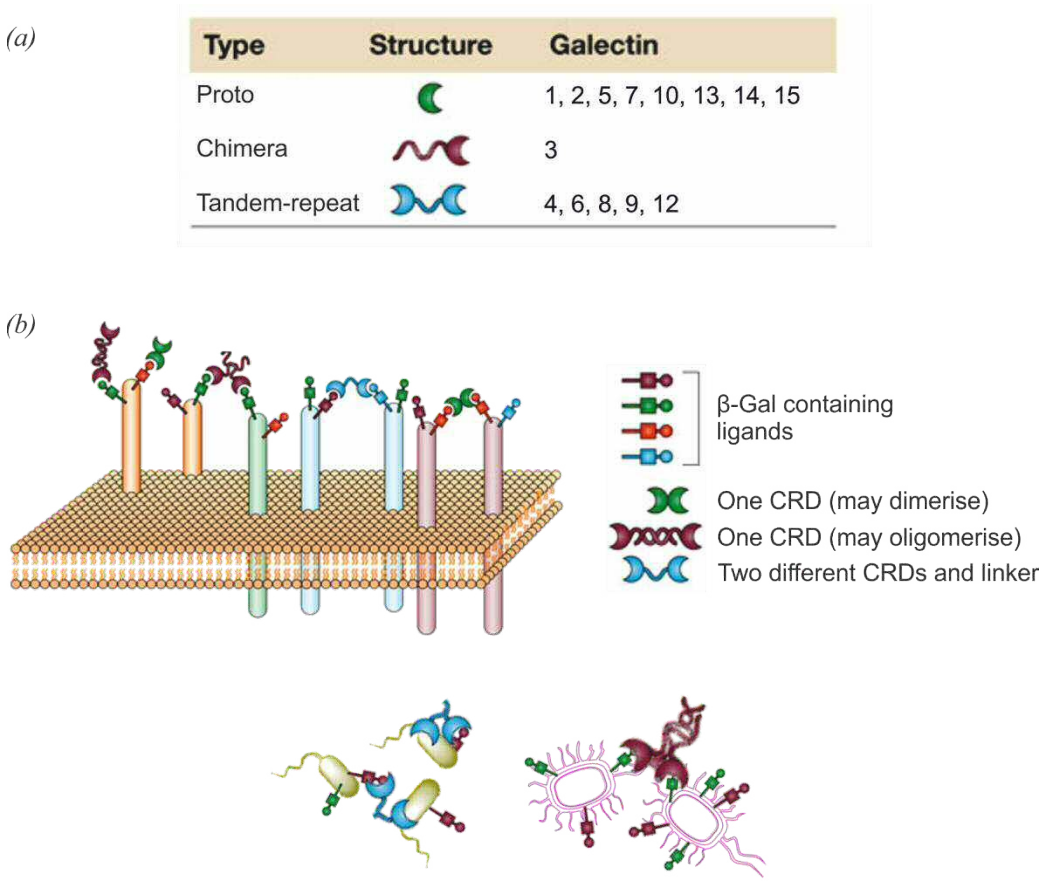


Fig. 4.26 Galectin classification and physiological roles. (*a*) Classification is based on CRD organisation (Hirabayashi and Kasai, 1993). (*b*) Among their functional roles, galectins bind in a *cis* or *trans* manner to Gal-containing ligands on bacterial surfaces, with different outcomes. Adapted from (Liu and Rabinovich, 2005).

Microarray screening of six different NTHi clinical isolates (section 4.1.8) revealed binding of hGal-8 to all tested strains, and evidenced for the first time the availability of docking sites for this galectin on the surface of NTHi strains other than NTHi2019 (Kalograiaki *et al.*, 2016). In order to explore whether other members of the galectin family are also capable to recognise the Gal-bearing NTHi glycosylation patterns (section 4.1.3), and if possible dissect the specific sugar epitopes involved in the recognition, we have examined the binding of a panel of 6 human galectins including members of the three structural subgroups to array-printed NTHi375. All the tested galectins were available in biotin-labelled form, thereby enabling inter-lectin comparison of their binding intensities.

Binding of proto-type hGal-1 and hGal-2 at the maximum concentration tested (2 μM) was negligible (Fig. 4.27). Direct binding to live bacteria at exponential growth phase (described in section 3.9.1) was not visible either (data not shown), confirming the microarray results. Therefore, hGal-1 was used as negative control in all subsequent experiments.

In contrast, the CRD of hGal-3 (designated as hGal-3tr, "tr" accounting for truncated) bound to array-printed NTHi375, revealing the presence of docking sites for this lectin

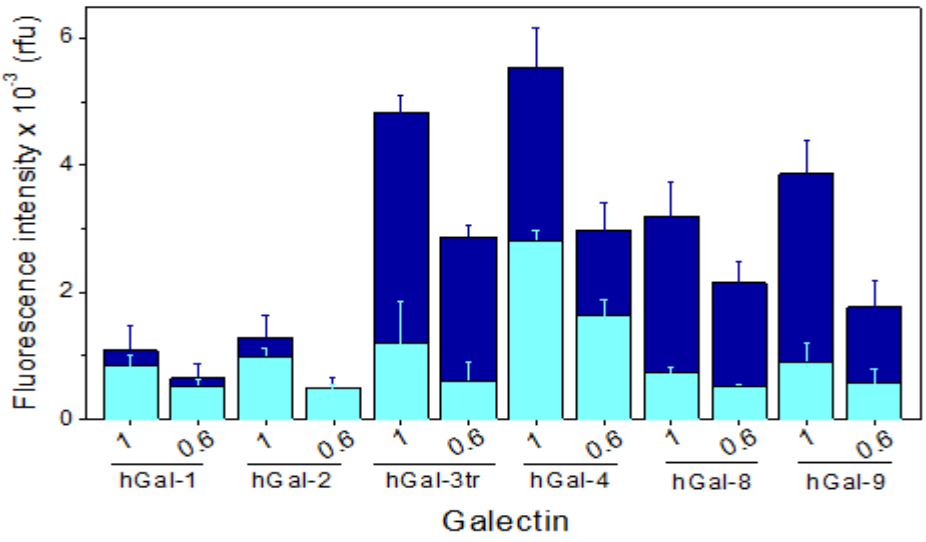


Fig. 4.27 Binding of biotinylated galectins to NTHi375 microarrays.
Binding of galectins in PBS$_\beta$ to NTHi375 printed as triplicates at OD_{600} = 1 and 0.6, was tested in the absence (blue) or presence of 0.1 M Lac (cyan). Galectin concentrations were set as follows: hGal-1 and -2: 2 μM, hGal-3tr: 1 μM, hGal-4: 0.25 μM, hGal-8 and -9: 0.5 μM. Data shown correspond to the mean of at least three independent experiments, upon normalisation of the amount of printed bacteria based on SYTO® 13 fluorescence signals and bacteria labelling efficiency. Error bars indicate the standard deviation of the mean.

on the pathogen surface, and the binding was significantly reduced in the presence of Lac, indicating carbohydrate-mediated recognition (Fig. 4.27).

On the other hand, binding of hGal-4 was less efficiently inhibited in presence of either 0.1 M Lac (Fig. 4.27), or 0.1 M LacNAc (~55% inhibition, not shown), the latter previously reported to be slightly better inhibitor for both hGal-4 CRDs (Buzamet, 2012). This might be due to the establishment of high-avidity interactions resulting from ligand clustering, and/or to non-carbohydrate-mediated secondary interactions of *e.g.* electrostatic or hydrophobic nature. In contrast, hGal-8 and hGal-9, the other two tandem-repeat galectins tested, bound to NTHi375 in a carbohydrate-specific manner (Fig. 4.27).

Having detected binding of the chimera and tandem-repeat type galectins, we next explored in more detail possible factors involved in the recognition, starting with the impact of the presence of the non-lectin N-terminal region of hGal-3 on the behaviour of this lectin.

4.1.10.1 NTHi375 recognition by hGal-3

SDS-PAGE analysis of bacterial lysates following incubation with non-labelled hGal-3 of NTHi375 suspensions in exponential growth phase showed the presence of a protein band with the electrophoretic mobility of the full length protein (Fig. 4.28, *a*), pointing to direct binding of hGal-3 to live bacteria. Densitometric quantitation of the band revealed a 50% decrease in intensity when the incubation was performed in the presence of 50 mM Lac, indicative of carbohydrate-mediated binding. Moreover, as observed by confocal microscopy, full length hGal-3 was able to agglutinate fixed bacterial cells (Fig. 4.28, *b*), what necessarily implies oligomerisation. Indeed, N-terminal tail-mediated pentamerization of hGal-3 in the presence of multivalent ligands has been reported (Ahmad *et al.*, 2004).

No effect of full length hGal-3 on NTHi adhesion to A549 cells was observed in *in vitro* assays (see section 3.9.3) posing the question of the physiological role of the hGal-3–NTHi interaction. Interestingly, hGal-3 also binds to the surface of *Neisseria meningitidis* (Quattroni et al., 2012), a frequent colonizer of the human respiratory tract. The binding requires intact LPS molecules and is coupled to an increase in adhesion to

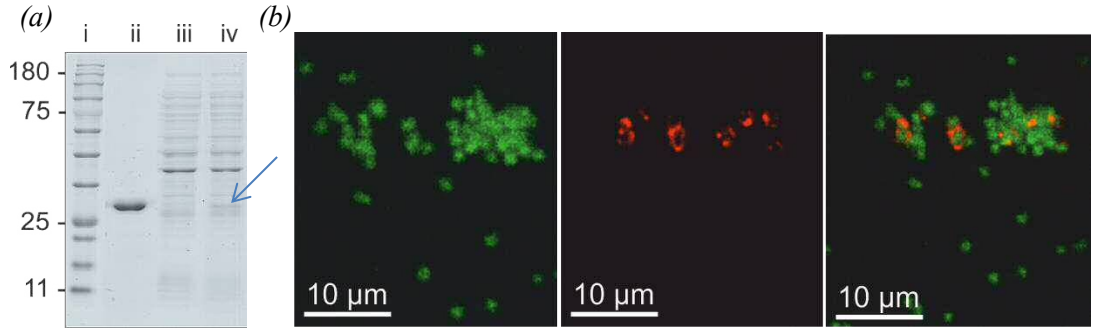


Fig. 4.28 Binding of hGal-3 to NTHi375. *(a)* Representative electrophoretic profiles of NTHi375 lysates in 12% polyacrylamide gels under reducing conditions. For the binding assay, 150 µl of a bacteria suspension (OD_{600} = 1) in PBS_{β} was used. (i) MW markers; (ii) 5 µg of non-labelled hGal-3; *bacterial pellet lysates:* (iii) control in the absence of galectin; (iv) bacteria incubated with 10 µM non-labelled hGal-3. *(b)* Binding of 10 µM His-tagged hGal-3 (central panel) to NTHi375 stained with SYTO® 13 (left panel) was detected by incubation with 1 µg/ml AF647-labelled streptavidin as final step (see 3.9.4). The right panel corresponds to the merge of SYTO® 13 and AF647 fluorescence signals.

monocytes and macrophages, apparently without contributing to phagocytosis, but not to epithelial cells. In contrast, hGal-3 has been found to significantly increase phagocytosis of *Klebsiella pneumoniae* (strain Kp52145) by human macrophages (unpublished results), thus indicating that the outcome of hGal-3 binding is bacterium specific. As *K. pneumoniae* LPS is also a ligand for rat Gal-3 (Mey *et al.*, 1996), the possibility that the NTHi LOS could serve as docking site for the lectin was considered. However, taking the binding to wild type NTHi375 as 100%, the relative binding strength in the bacteria arrays to entire NTHi375Δ*ompP5* was 98 ± 5% and 108 ± 6% for full-length hGal-3 and its separate CRD, respectively, indicating that LOS overexpression in the Δ*ompP5* mutant does not result in increased binding. Furthermore, no binding of either hGal-3 or hGal-3tr to array-printed NTHi375Δ*ompP5* LOS was observed. Altogether, the results argue against the possibility that the NTHi375 LOS behaves as ligand for hGal-3.

4.1.10.2 NTHi375 recognition by tandem-repeat galectins hGal-4, hGal-8, and hGal-9

Binding assays to the surface of bacteria at exponential growth phase, performed as described above for hGal-3, revealed the presence of protein bands with the electrophoretic mobility of full length hGal-4 (Fig. 4.29, *a*) and hGal-8 (Fig. 4.29, *b*). The intensity of the hGal-8 band decreased 20-30% in the presence of 50 mM Lac, indicating that binding to live bacteria is, at least in part, carbohydrate-mediated.

However, no inhibition at all was observed for hGal-4 under the assayed conditions, here pointing to the establishment of non-carbohydrate-mediated interactions. Regarding hGal-9, the appearance of protein bands at the position expected for a separate CRD evidenced degradation most probably due to linker proteolysis (Fig. 4.29, *c*), a phenomenon frequently observed for tandem-repeat-type galectins. Still, the lectin proved to be functional, as a protein band at the position of full length hGal-9 was clearly visible in the bacterial lysates following incubation with this galectin, and the intensity of the band decreased 50-55% in the presence of 50 mM Lac.

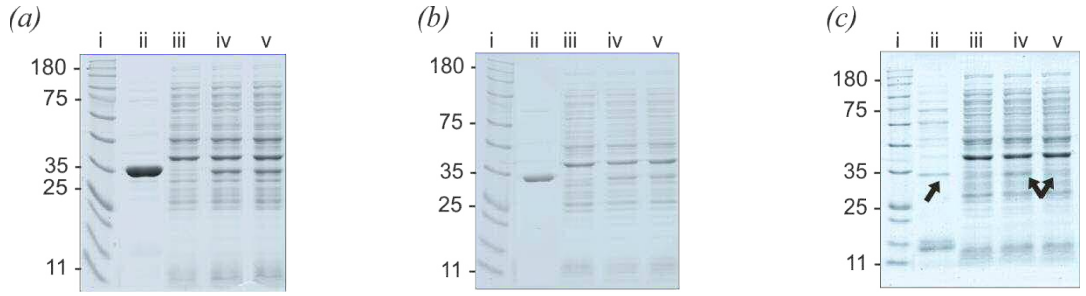


Fig. 4.29 Direct binding of tandem-repeat hGal-4, hGal-8 and hGal-9 to live NTHi375. Representative electrophoretic profiles in 12% polyacrylamide gels under reducing conditions of NTHi375 lysates after incubation of bacteria suspensions (OD_{600} = 1) in PBS_{β} with 10 µM non-labelled hGal-4 (*a*), hGal-8 (*b*), or hGal-9 (*c*), in the absence (lanes iv) and presence of 50 mM Lac (lanes v). (i) MW markers; (ii) 5 µg of non-labelled galectin; (iii) NTHi375 control pellet lysate.

In contrast to hGal-3, the binding of the three tandem-repeat galectins increased considerably as result of LOS overexpression, with relative binding strengths to array-printed NTHi375Δ*ompP5* of ~250 ± 20% for hGal-4 and -9, and 170 ± 10% for hGal-8 with respect to wild type NTHi375. Thus, in this case the LOS appears a likely candidate to act as ligand for these lectins. Therefore, the binding of hGal-4, -8 and -9 to the panel of mutants presenting sequentially truncated LOSs was next examined (Fig. 4.30). The three galectins exhibited reduced binding to the Δ*siaB* mutant whereas silencing of *lgtF*, the gene coding for the glycosyltransferase responsible for adding Glc to Hep I, had no significant effect or even increased binding, a behaviour also observed for the anti-NTHi serum (section 4.1.1) and the plant lectins VAA, SBA and ConA (section 4.1.3). Truncations affecting the extension at Hep III (*i.e.* Δ*lic2A*, Δ*lpsA*, and Δ*lgtF*Δ*lpsA*) or even all core sugars (Δ*opsX*) had only small consequences on NTHi recognition by hGal-4 and hGal-9. However, they did weaken significantly the binding of hGal-8, for the Δ*opsX* mutant the relative binding strength being reduced up to 37 ±

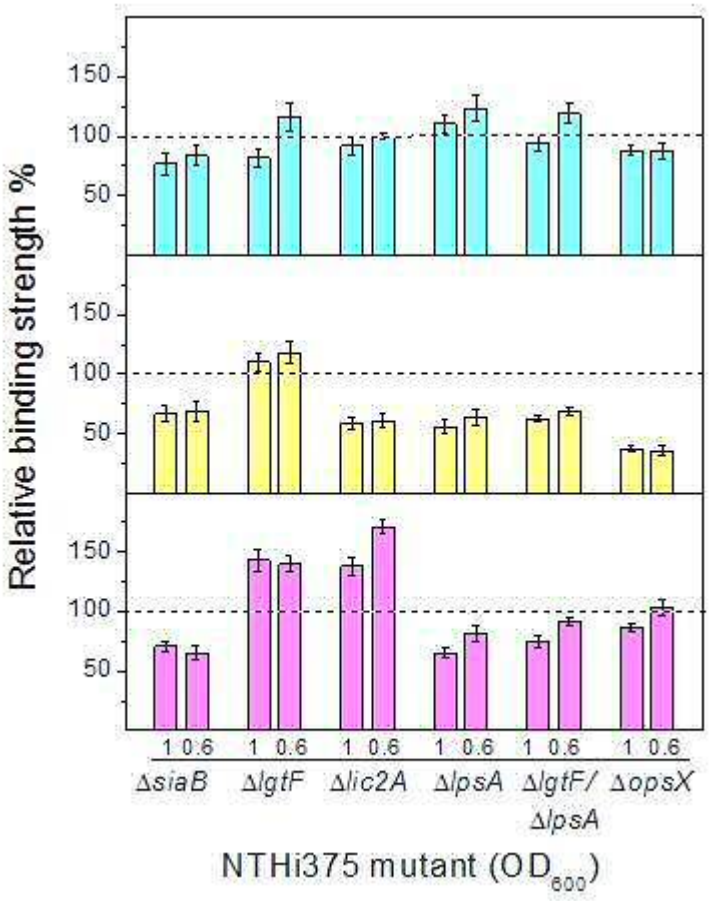


Fig. 4.30 Relative binding strength of hGal-4, hGal-8, and hGal-9 to NTHi375 mutant strains. Binding of 0.5 µM hGal-4 (cyan), hGal-8 (yellow) and hGal-9 (magenta) in PBS_{β} to NTHi375 wild type and mutant strains, printed as triplicates of OD_{600} values 1 and 0.6, was assessed simultaneously, and the relative binding strength was calculated as percentage taking binding to wild type NTHi375 as 100%. Data shown correspond to the mean of at least three independent experiments and error bars indicate the standard error.

1% with respect to wild type NTHi375. Thus, the Hep III extension could serve as docking point for hGal-8.

A distinguishing feature of tandem-repeat galectins is the presence of two different CRDs within the same molecule, enabling these lectins to recognize different sets of ligands. In order to get information on the relative contribution of the N- and C-terminal CRDs to NTHi375 binding, we took advantage of the availability of the separate domains of hGal-4 and hGal-8 (designated as hGal-4/8 N and C), which have been shown to exhibit clearly distinct fine-tuned specificities (Huflejt and Leffler, 2003; Carlsson *et al.*, 2007).

Using the same strategy followed for the full-length galectins, we first examined the binding of the domains to live NTHi375 in exponential growth phase. As detailed in Table 4.10, no significant binding of hGal-4N was detected. In addition, the intensity of

Table 4.10 Binding of hGal-4 and -8 CRDs to NTHi375 at exponential growth phase.

CRD	Optical density ($\times 10^{-2}$)[a]		
	Control pellet	Pellet + galectin	Pellet + galectin + 50 mM Lac
hGal-4N	6 ± 1	7 ± 1	6 ± 1
hGal-4C	8 ± 1	26 ± 4	28 ± 5
hGal-8N	6 ± 1	36 ± 4	19 ± 3
hGal-8C	6 ± 1	11 ± 1	9 ± 2

Measured by densitometric quantitation of the CRD protein bands upon SDS-PAGE analysis of bacterial lysates following incubation of NTHi375 suspensions in exponential growth phase with 10 μM of the respective CRD, in the absence or presence of 50 mM Lac. The intensity of the control pellet corresponds to the background at the equivalent electrophoretic mobility. [a]Intensity value for the center of the band (3x3). Data correspond to the mean of at least two different experiments and standard error values are given.

the hGal-4C and hGal-8C bands appearing in NTHi375 lysates following incubation with these two proteins did not decrease in the presence of 50 mM Lac. Only for hGal-8N, a noticeable inhibition of the binding in the presence of the hapten was observed, here evidencing carbohydrate-mediated recognition.

To get further insights into the possible ligands recognized by the hGal-4 and -8 separate CRDs, their binding to the panel of NTHi375 mutants was next examined using the bacteria microarray set-up (Fig. 4.31). The first remarkable observation is that binding of hGal-4C but not of hGal-4N was detected, in full agreement with the assays using live bacteria. Recognition by the C-domain was strongly favoured for the Δ*ompP5* mutant. However, truncation of the LOS had dissimilar consequences depending on the mutant, the binding to the shortest Δ*opsX* variant being comparable to that of the wild type strain.

Regarding hGal-8, although binding of both CRDs was observed, the consequences of LOS overexpression and truncation were entirely opposite. Thus, for hGal-8C no increase in the binding was found for the Δ*ompP5* mutant compared to wild type NTHi375 and LOS truncation at certain levels even favoured recognition. In striking contrast, recognition by the N-domain was strongly favoured for the Δ*ompP5* mutant, as

Lectin	Mutant strain						
	Δ*siaB*	Δ*lgtF*	Δ*lic2A*	Δ*lpsA*	Δ*lgtF* Δ*lpsA*	Δ*opsX*	Δ*ompP5*
hGal-4N	NO BINDING						
hGal-4C	81 ± 5	76 ± 4	78 ± 6	49 ± 4	60 ± 6	92 ± 5	250 ± 17
hGal-8N	58 ± 6	89 ± 4	60 ± 5	41 ± 7	58 ± 8	68 ± 6	210 ± 6
hGal-8C	67 ± 3	150 ± 12	102 ± 10	131 ± 9	140 ± 10	99 ± 6	89 ± 6

> 150 %	150 – 115	115 – 85	85 – 50	50 – 25	< 25 %

Fig. 4.31 Relative binding strength of hGal-4 and -8 separate CRDs to NTHi375 microarrays. Binding of 0.5-5 µM CRDs in PBS_{β} to immobilised NTHi375 wild type and mutant strains, printed as triplicates at OD_{600} values 1 and 0.6, was assessed simultaneously, and the relative binding strength was calculated as percentage taking binding to wild type NTHi375 as 100%. Data shown correspond to the mean of at least 2 independent experiments and the standard error is indicated.

similarly observed for hGal-4C, and truncation of the LOS extension at Hep III resulted in decreased binding (Fig. 4.31), suggesting that this branch could serve as docking point for hGal-8N. Only silencing of *siaB* had a similar impact on the two hGal-8 CRDs, with a noticeable drop in binding. Altogether, the results evidence that the N- and C-terminal CRDs of hGal-8 recognize different ligands on the NTHi375 surface.

Indeed, different roles have been proposed for the two domains of tandem-repeat galectins when interacting with Gram-negative bacteria, one CRD serving for carbohydrate recognition and the other contributing to a loss of membrane integrity, leading to bacteria cleavage (Stowell *et al.*, 2010), susceptibility of NTHi2019 to hGal-8 having been reported (Stowell *et al.*, 2014). Thus, the contrasting behaviour of the hGal-8N and C- CRDs in the microarray binding assays could be reflecting these different roles.

Based on the important increase in the binding of hGal-4C and hGal-8N to NTHi375 as a result of LOS overexpression, and on the impact of LOS truncation, the LOS appears to act as ligand for these CRDs. Therefore, their binding to array-printed LOSs was examined in comparison with the full-length galectins (Fig. 4.32). An important observation is that the behaviour of hGal-4 and hGal-8 paralleled that of their respective C- and N-domains, indicating that these CRDs are mainly responsible for the binding of the full-length proteins to the tested LOSs.

For hGal-8, a stronger binding to Δ*lgtF* over Δ*ompP5* LOS was observed, but truncation of the LOS extension at Hep III in Δ*lpsA* dramatically reduced recognition (Fig. 4.32),

evidencing that this branch behaves as ligand for hGal-8N. Moreover, in the presence of 0.1 M Lac a substantial inhibition of the binding to Δ*ompP5* (70-80%) and Δ*lgtF* (90-92%) LOS was observed for both the full-length protein and N-domain, inhibition reaching values of up to 97% in the presence of 0.01 M 3´-sialyl-lactose, known to be a superior ligand for hGal-8N (Ruiz *et al.*, 2014).

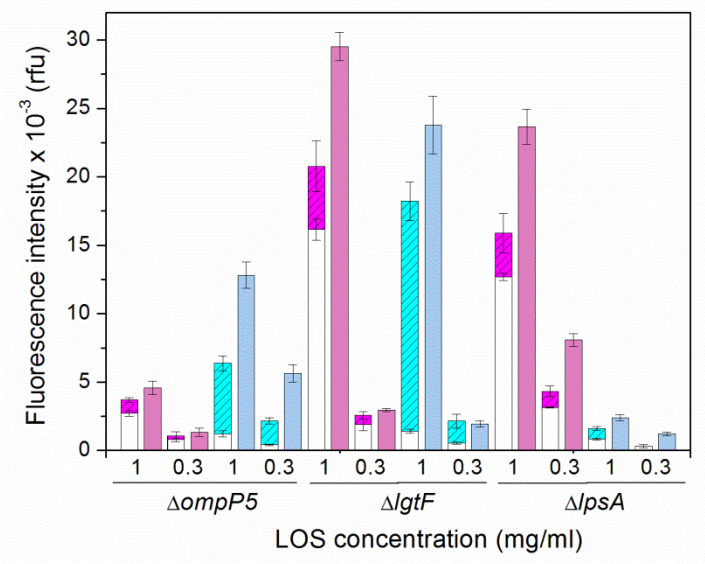


Fig. 4.32 Binding of hGal-4 and hGal-8 to NTHi375 LOSs. The LOSs were immobilised as triplicates at decreasing concentrations and the binding of 0.5 µM hGal-4 (magenta), hGal-4C (light magenta), hGal-8 (cyan) or hGal-8N (light cyan) in PBS_{DTT} was examined. Competition experiments were run with the full-length galectins in the presence of 0.1 M Lac (white). Only the two highest concentrations of printed LOSs are shown. Error bars stand for the standard deviation of the mean.

An entirely different picture was obtained for hGal-4 (Fig. 4.32). Here, a noticeably higher binding to both Δ*lgtF* and Δ*lpsA* LOSs over the Δ*ompP5* LOS was detected. However, in all cases inhibition was hardly observed in the presence of Lac (Fig. 4.32), LacNAc, or its 3´/6´-sulphated derivatives (not shown). Contrarily, the binding was completely abolished by high ionic strength (1 M NaCl), indicating the involvement of electrostatic interactions. What is more, hGal-4 strongly bound to the isolated lipid A of NTHi375, binding being partially inhibited (~30%) in presence of 1'/6'-phospho-GlcNAc, and to a greater extent (60-90%, depending on lipid A concentration) by 1 M NaCl (not shown). Taken together, the results argue for ionic, non-carbohydrate-mediated interactions being the main forces driving NTHi375LOS recognition by hGal-4. A further distinctive feature of hGal-4 was a significant increase in binding to both NTHi375 LOS and lipid A when the lectin was assayed in the absence of reducing agents (DTT or *β*-ME). Indeed, confocal microscopy demonstrated a dramatic increase

in bacteria agglutination under non-reducing conditions (Fig. 4.33). The rationale behind this behaviour will be the subject of future studies.

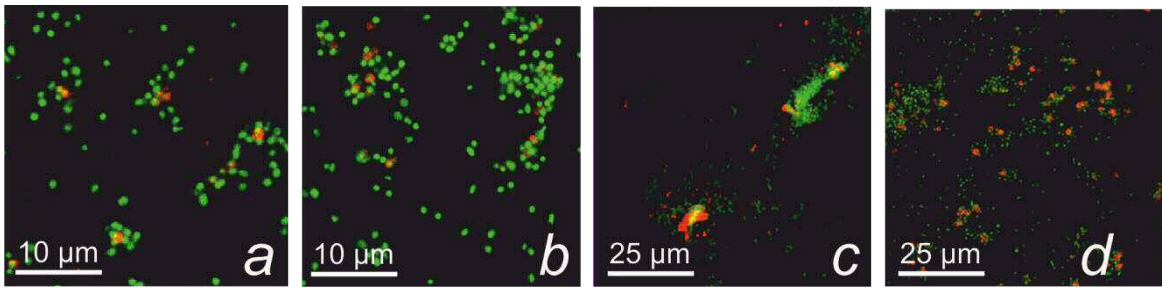


Fig. 4.33 Effect of red-ox conditions on hGal-4 binding to NTHi375.
(*a*, *b*) hGal-4 binding (10 µM, red) to NTHi375 stained with SYTO® 13 (green) in the presence of 1 mM DTT. (*c*, *d*) hGal-4 binding (10 µM) in absence of reducing agent. In all cases, experiments were run in duplicates, pictures representing different preparations.

At any rate, no effect of hGal-4 on the colony formation capacity of NTHi375, or on its adhesion to or invasion of pulmonary epithelium (cell line A549) was observed either in the absence or presence of reducing agents.

4.2 Exploring glycosylation profiles of extracellular vesicles

Our understanding of the mechanisms of intercellular communication and transport increased with the discovery that sub-micrometric membranous particles can act as carriers for "message exchange" (György *et al.*, 2011b; Mittelbrunn and Sánchez-Madrid, 2012). Extracellular vesicles (EV), encompassing exosomes (EXO), microvesicles (MV), and apoptotic bodies, are bilayer-enclosed assemblies originated from most mammalian cell types (Raposo and Stoorvogel, 2013), and isolated from conditioned cell culture media or biological fluids (blood, urine, breast milk, etc.) (Boukouris and Mathivanan, 2015).

A clear-cut dissection of EVs' composition and functional roles in health and disease is currently hampered by the diversity of approaches used for their characterisation, making the establishment of common guidelines essential (Lötvall *et al.*, 2014). Isolation methodologies usually used, primarily based on differential centrifugation or density gradient separation, also complicate obtaining pure EV fractions. Consequently, the composition of EV preparations greatly varies depending on the protocol used. In addition, information on their long-term stability is limited (Kalra *et al.*, 2013; Zonneveld *et al.*, 2014).

Physical characterisation of EV preparations at least requires observation of size, structure and morphology, while their concentration is expressed in terms of particle number, mass or, gradually more frequently, their biomolecular cargo (Rupert *et al.*, 2017). Aiming at retrieving data from proteomic, lipidomic and transcriptomic analyses (Choi *et al.*, 2013), several web-based databases have been launched, as *e.g.* Vesiclepedia (Kalra *et al.*, 2012), EVpedia (Kim *et al.*, 2013), and ExoCarta (Simpson *et al.*, 2012). The latter is a key resource for exosomal cargo (Keerthikumar *et al.*, 2016), unfortunately only offering comprehensive data for this EV subpopulation.

Still, descriptive studies on the dynamic glycan profile of distinct subpopulations from different sources are currently missing. In this context, only a limited number of reports have been published (Krishnamoorthy *et al.*, 2009; Gomes *et al.*, 2015), along with several studies suggesting a role in the MV internalisation by target cells (Escrevente *et al.*, 2011), or focussing on glycan-mediated sorting of cargo into MV (Batista *et al.*, 2011; Liang *et al.*, 2014), and on the search of cancer biomarkers (Melo *et al.*, 2015).

Therefore, in this Thesis, we have aimed at developing a widely applicable analytical methodology for glyco-profiling distinct EXO and MV populations. The THP-1 human acute monocytic lymphoma cell line was selected as EV source. Characterisation of these populations have included determination of total protein cargo and lipid content, particle concentration by resistive pulse sensing (RPS), and morphology observation by transmission electron microscopy (TEM). Then, following immobilisation as microarray probes, recognition of EVs by our panel of lectins with diverse carbohydrate-specificities was examined. In addition, lectin-derivatised QCM chips were developed to enable "fishing" and detection of EXO in flow. Finally, binding kinetics for representative EV–lectin pairs was investigated using EV-based QCM chips.

4.2.1 Characterisation of THP-1-derived EXO and MV populations

Quantitation and normalisation of EV preparations, an essential step prior to downstream assays, can be approached by determination of their total protein content. However, its use as single reference is devalued by the likely presence in the preparations of protein complexes and/or aggregate contaminants, which are ubiquitous in body fluids and culture media (György *et al.*, 2011a). The "total" lipid determination using sulfuric acid and vanillin complement protein cargo measurements, and protein-to-lipid ratios have recently proved to be suitable for distinguishing among EV subpopulations (Osteikoetxea *et al.*, 2015).

In this Thesis, the protein-to-lipid ratio and RPS measurements (section 3.3.2) were used for determination of EV concentration (Table 4.10). For all the preparations tested, EXO gave rather uniform size measurements of 90-110 nm (mean), whereas MV were highly heterogeneous, with a calculated mean of 220-270 nm and the occurrence of larger particles (Fig. 4.34, *a*). For selected preparations, TEM pictures were obtained (Fig. 4.34, *b*, for experimental details see section 3.3.3). Data for all preparations are given in Table 4.11.

Membrane lipid order has also been reported to be a distinctive feature of different EV subpopulations (Osteikoetxea *et al.*, 2015). An indirect yet easy way to infer the fluidity of membranes is to use polarity sensitive probes whose emission spectra change with

Table 4.11 Protein cargo, protein/lipid ratio and particle quantitation measurements for THP-1-derived EV preparations

	EXO				MV			
	Protein (µg/ml)	**Lipid (µg/ml)**	**Protein-to-lipid**	**RPS (U/ml)**	**Protein (µg/ml)**	**Lipid (µg/ml)**	**Protein to-lipid**	**RPS (U/ml)**
THP-1a	0.2	0.85	**0.2**	9.1×10^{11}	4.1	6.8	**0.6**	2.6×10^{11}
THP-1b	13.05	59.7	**0.2**	5.2×10^{13}	4.2	2.3	**1.8***	1.4×10^{11}
THP-1c	0.9	1.00	**0.9***	ND	0.9	0.8	**1.2***	ND
THP-1d	4.8	17.5	**0.27**	3.1×10^{12}	3.1	5.7	**0.5**	ND

EVs were isolated from a THP-1 cell culture (in suspension) and purified as described in section 3.2. Lipid concentration was determined at the laboratory of Prof. E. Buzas (Semmelweis University, Budapest). RPS values correspond to stock preparations. The asterisk indicates protein contamination as confirmed by TEM (see Fig. 4.34, *b*). ND, not determined.

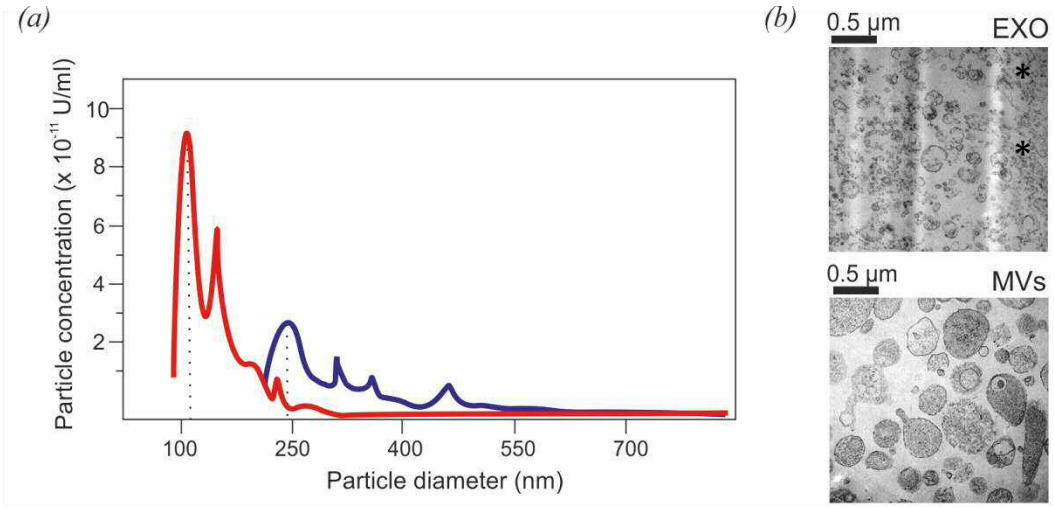


Fig. 4.34 Characterisation of the size and morphology of THP-1-derived EV.
(a) Representative RPS histograms for EXO (red) and MV (blue) of preparation THP-1. At least 500 measurements were recorded for serial dilutions of EV preparations in PBS and particle concentrations calculated from each dilution for the stock preparation were merged into a single histogram for each subpopulation. Major peaks correspond to the mean size of particles and are representative of particle concentration (also see Table 4.11). *(b)* TEM images of EV preparations derived from THP-1 culture of advanced confluency and overnight vesiculation in serum depleted media (section 3.3.1). Vesicles were fixed in paraformaldehyde, post-fixed with OsO4, and the resin block was stained with uranyl-acetate (section 3.3.3). The asterisks spot protein contamination.

the polarity of the environment (Parasassi *et al.*, 1998). A ratiometric determination of general polarisation (GP) values was carried out by fluorescence spectroscopy, upon staining with the polarity-sensitive lipid probe di-4-ANEPPDHQ (section 3.3.4). The GP value was calculated from the difference in fluorescence intensity at 570 nm and 630 nm (Fig. 4.35), resulting in values between −1 and +1, the higher the GP value the higher the lipid order (*i.e.* low liquid disorder and more packed membrane structure),. THP-1-derived EXO displayed GP values of 0.37-0.39, reflecting a remarkably higher

lipid order than MV, which displayed GP values of 0.07-0.1 (based on measurements on three different preparations of each subpopulation). These values were comparable to those previously reported (Osteikoetxea *et al.*, 2015).

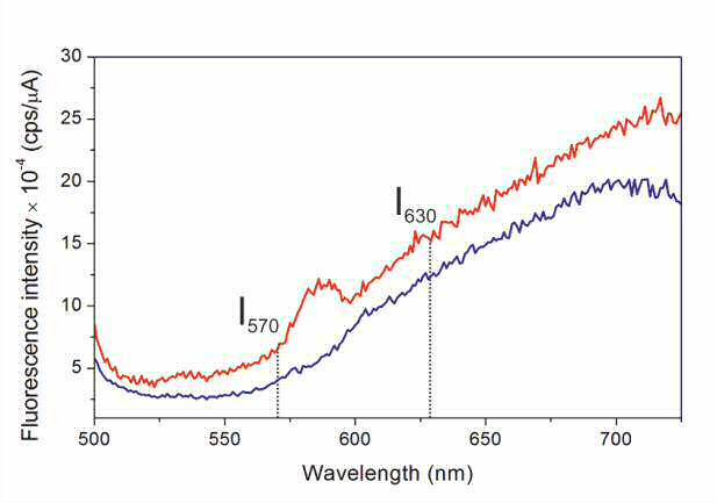


Fig. 4.35 Ratiometric determination of EV membrane lipid order by fluorescence spectroscopy. Representative emission spectra of THP-1-derived EXO (red) and MV (blue) preparations stained with di-4-ANEPPDHQ, upon excitation at 488 nm. Intensities corresponding to membrane ordered (570 nm) or disordered (630 nm) regions were used for ratiometric determination, as described in section 3.3.4.

4.2.2 Preparation and validation of EV microarrays

EV were printed onto nitrocellulose-coated glass slides as triplicates at increasing dilutions. Only preparations with a meaningful EV concentration and confirmed absence of protein contamination, as derived from the protein-to-lipid ratio and TEM measurements, were used. EV could not be fluorescently labelled following the approach previously described for bacteria (using thenucleic acid stain SYTO-13®) due to their random and rather low content in nucleic acids, yielding signals far below the detection limit of the microarray scanner. On the other hand, the lipid-packing indicator di-4-ANEPPDHQ emits in the "red" region upon excitation, thus interfering with our detection system based on the use of AlexaFluor®647. Therefore, for validation of the developed EV-microarrays the binding of specific anti-tetraspanin antibodies and lipid-binding molecules (annexin V) was examined. Tetraspanins (CD63, CD9, etc.) have been proposed as exosomal markers due to their enrichment in this vesicle subpopulation, from 7- to 124-fold compared to the parental cells (Andreu and Yáñez-Mó, 2014). They are also found on the plasma membrane of certain cell-types hinting at their presence also on MV. In addition, annexin V conjugates are conventionally used for MV detection, as they bind phosphatidylserine, which is "flipped" and exposed on the plasma membrane during vesiculation (Dachary-Prigent *et al.*, 1993).

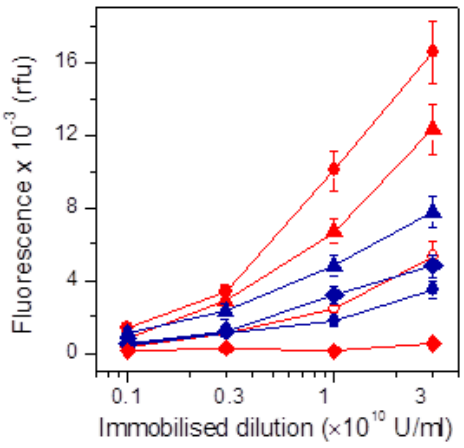


Fig. 4.36 Validation of EV microarrays. NC-printed THP-1- (triangles) and blood-derived (circles) EXO (red) and MV (blue) were incubated with anti-CD63 at 1:1000 dilution and binding was detected by use of a secondary biotinylated antibody and streptavidin-AF647 (circles for Alternatively, they were incubated with Annexin V-FITC (1:100) in Ca^{2+}-containing buffer (diamonds). Open circles correspond to THP-1-derived EXO supernatants at the same dilution used for EV.. Data shown correspond to the mean of two different experiments and error bars indicate the standard deviation.

Both array-printed THP-1-derived EXO and MV populations were reliably detected by anti-CD63 antibody, the binding showing the expected concentration-dependent trend (Fig. 4.36). Surprisingly, moderate binding signals to printed supernatants from EXO ultracentrifugation were detected, pointing to the presence of small size vesicles that may be lost using the conventional centrifugation protocol, as similarly noticed in recent studies on urinary extracellular vesicles (Musante *et al.*, 2017). Tested in parallel, annexin V-FITC did not bind THP-1-derived MV even upon buffer supplementation with 10 mM Ca^{2+}, in line with previous observations (Ismail *et al.*, 2013). However, it selectively recognised blood-derived MV, what could result from the use of human blood without prior platelet depletion as starting source for EV isolation (Thiagarajan and Tait, 1990).

In any case, both anti-CD63 antibody and annexin V bound to the immobilised vesicles in a dose-dependent manner, proving accessibility of vesicle surface epitopes and opening the way to EV glycoprofiling.

4.2.3 Profiling of EV glycosignatures

Screening of our panel of 19 biotinylated lectins of known binding specificities (see Table 3.2 for details) revealed noticeable differences in binding depending on the EV subpopulation (Fig 4.37). For all of the lectins, the specificity of the recognition was assessed in parallel assays carried out in the presence of their respective inhibitory hapten.

Except for HPA and DSL, moderate to strong binding to THP-1-derived EXO and/or MV were observed for all the lectins, the Gal-specific VAA, RCA, and MAL-I

agglutinins giving intense signals for both EV subpopulations. In general, similar lectin-binding trends were observed, with the following significant exceptions. First, divergent binding intensities to EXO and MV were found when comparing PNA (exhibiting preference for Galβ(1,3)GalNAc) and WFA (specific for GalNAcα(1,3/6)Gal), and the same was found for HHL (preferentially recognizing Manα(1,3/6)Man) *vs* PSA (which shows a broad binding specificity towards Man, Glc, and GlcNAc), revealing a different availability of the recognized epitopes on the EV surfaces. Second, WGA (selective for GlcNAc and Neu5Ac) binding signals for MV almost doubled those for EXO, while the opposite was found for PSqL (specific for Neu5Acα(2,6)Galβ(1,4)GlcNAc/Glc). And third, a strikingly EV-selective behaviour of GNA (recognizing terminal Manα(1,3)Man) was observed, this lectin giving the strongest binding signals for EXO

populations and only very moderate signals for MV. The intense binding of ConA (Man/Glc specific) to EXO further supported a higher abundance of Man-containing structures on the EXO surface. Of note, GNA and ConA binding was only partially inhibited in the presence of 0.1 M α-Me-Man (Fig. 4.37), whereas 10 mg/ml mannan from *S. cerevisae* was able to reduce binding below 15% (not shown), indicating the establishment of high avidity GNA/ConA–EXO interactions.

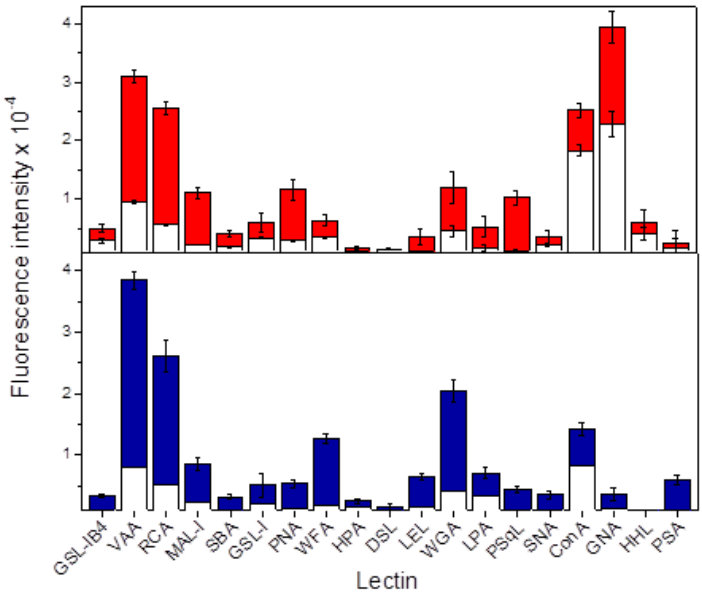


Fig. 4.37 Binding of reference lectins to THP-1-derived EXO and MV
EXO (upper panel) and MV (lower panel) preparations were printed in triplicate at 0.1-3 × 10^{10} U/ml and the binding of biotinylated lectins in the absence (red/blue) and presence (white) of the respective lectin hapten (Table 3.2).was assessed by incubation with AF647-streptavidin. Data shown correspond to the mean of at least two experiments, each performed with different EV preparations. Error bars indicate the standard deviation of the mean.

In the case of blood-derived EV, strong binding signals for several lectins were also detected, although no clear binding trends were observed (not shown). This behavior could be due to the diversity of cells contributing to EV occurring in blood. Still, GNA exhibited a clear preference for EXO populations, arguing in favour of a highly-specific EXO glycosignature.

Altogether, the results unveil differential glycosylation patterns for these two vesicle subpopulations that could be exploited for designing new isolation protocols based on lectin affinity chromatography.

4.2.4 QCM-assisted detection of THP-1-derived EV under continuous flow

Based on the microarray results suggesting specific recognition of THP-1 EXO populations by GNA, in-flow detection by a QCM biosensor was targeted. To this end, GNA-derivatised chips were prepared as detailed in section 3.6.1. GNA immobilisation through amine coupling yielded frequency shifts somewhat smaller than those observed for ConA (50 vs 70 Hz, mean values respectively), pointing to a poorer surface coverage. Nevertheless, the chips proved to be efficient EXO capturers.

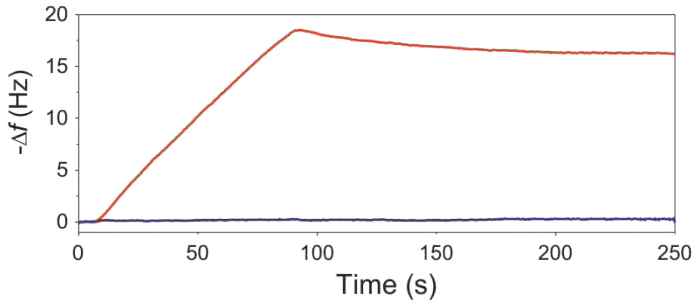


Fig. 4.38 Vesicle capture in flow by GNA-derivatised QCM chips.
1:10 dilutions in PBS of THP-1-derived EXO (red) and MV (blue) were injected at 20 µl/min and 22 ºC over a GNA-derivatised QCM chip (section 3.6.1). Samples were injected in duplicate, the mean sensor response being here presented.

For THP-1-derived EXO, remarkable frequency shifts were recorded, and the sensor response reached the steady state in a reasonable time upon injection (100 s) (Fig. 4.38). In contrast, THP-1-derived MV did not bind to the immobilised lectin, confirming the microarray results. The experiments were carried out with two different THP-1 EV preparations, showing good reproducibility, within the expected error limits for the estimated particle concentration (not shown).

Sample scarcity precluded a full kinetic analysis of the interaction, and equilibrium analysis for K_D calculation based on a single injection was not attempted, since the bound fraction may differ depending on the injected concentration and volume (Hunter and Cochran, 2016). Instead, since GNA was a confirmed capturer for EXO, we focused on the development of EXO-based sensor chips for the kinetic analysis of EXO-lectin interactions.

4.2.5 QCM EXO-based biosensors: monitoring lectin-binding kinetics directly at the vesicle surface

The concept and development of QCM EXO-based chips is described in detail in section 3.6.1, grounded on our previous observations with bacteria-based chips (Kalograiaki *et al.*, 2016). EXO chips enabled real-time monitoring of the binding of selected lectins to the vesicle surface and, importantly, allowed the reliable calculation of rate constants without the need for reaching equilibrium. Efficient regeneration by a single injection of glycine pH 2.5 was also an important advance, preserving the vesicle surface while avoiding the use of buffers with ionic strength beyond the physiological values. To be mentioned, following EXO capture, chips were fixed with formaldehyde so that the vesicles become an integrated part of the sensor chip thereby allowing the use of efficient regeneration solutions for completely removing bound lectins between successive binding assays.

Based on the microarray results revealing a distinctive behaviour of ConA, GNA, and PNA towards EXO *vs* MV, the interaction of these three lectins was examined. Examples of the sensograms obtained upon injection of the lectins in a range of increasing concentrations are shown in Fig. 4.39, unveiling lectin-specific responses. Comparison of the maximum frequency shift induced by injection of each lectin, after normalisation for the molecular mass, was consistent with a higher amount of GNA bound to the chip compared to ConA and PNA (Table 4.12), in full agreement with the binding trend observed in the microarray assays. For calculation of the kinetic parameters, experimental data were fitted using 1:1 or 1:2 binding models (Fig. 4.39). The first model was found suitable for fitting the binding profiles of ConA and PNA despite the signal noise. The dissociation rates for the two lectins were similar, but association was 10-fold faster for ConA, resulting in a proportionally higher binding affinity (Table 4.12). On the other hand, a 1:2 model was required for fitting the experimental data obtained for GNA, the difference in affinity between the two sets of

binding sites on the EXO surface primarily arising from a noticeably faster dissociation rate from the low-affinity sites. Overall, the dissociation constants calculated from the kinetic parameters differ in a 15-fold factor the maximum, being in the low micro-/submicromolar range (Table 4.12).

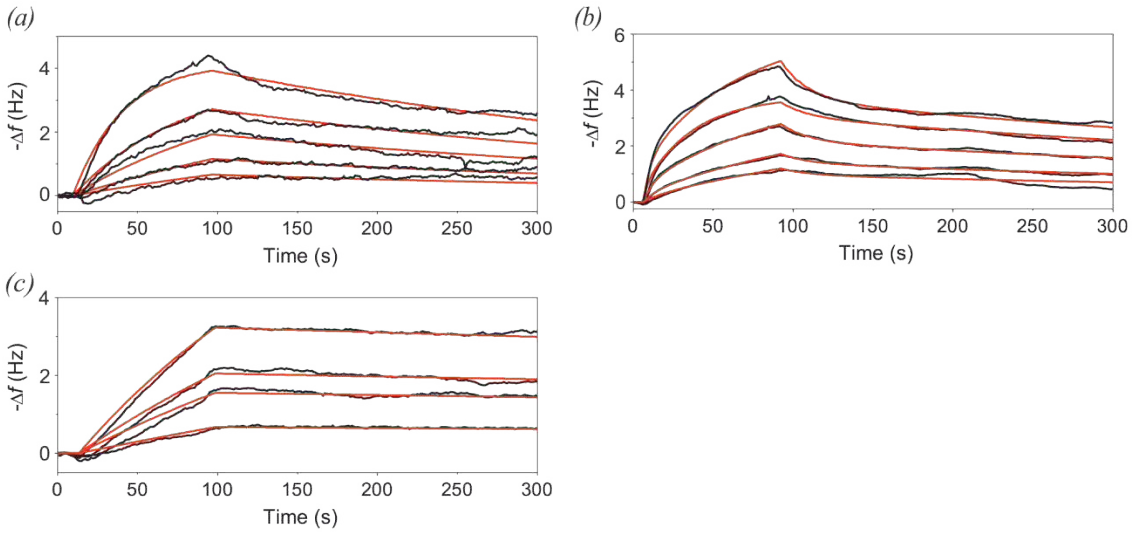


Fig. 4.39 Representative sensograms illustrating the binding of ConA, GNA and PNA to THP-1-derived EXO. Lectins were injected at the following concentrations: *(a)* 10.6, 35, 70, 104 and 175 µg/ml for ConA; *(b)* 34, 52, 86, 104 and 156 µg/ml for GNA; *(c)* 73, 181, 220 and 330 µg/ml for PNA. The black lines in the sensogram represent experimental data, and the red lines the fits obtained using 1:1 (ConA and PNA) or 1:2 (GNA) models.

Table 4.12 Analysis of the binding of the three lectins to EXO-derivatised QCM sensor chips.

	k_a (M^{-1} s^{-1})	k_d (s^{-1})	K_D (M)	$-\Delta f$ (Hz)	Normalised $-\Delta f$ (Hz/kDa)
ConA	$1.81 \pm 0.03 \times 10^4$	$2.5 \pm 0.1 \times 10^{-3}$	$1.4 \pm 0.1 \times 10^{-7}$	2.2	0.02
GNA	$5.08 \pm 0.05 \times 10^3$ $6.1 \pm 0.7 \times 10^4$	$1.53 \pm 0.01 \times 10^{-3}$ $5.76 \pm 0.07 \times 10^{-2}$	$3.01 \pm 0.04 \times 10^{-7}$ $9.4 \pm 0.7 \times 10^{-7}$	2.4	0.05
PNA	$1.8 \pm 0.1 \times 10^3$	$3.7 \pm 0.8 \times 10^{-3}$	$2.06 \pm 0.7 \times 10^{-6}$	1.7	0.02

The association rate (k_a), dissociation rate (k_d) and dissociation constant (K_D) are derived from the fitting of the experimental data to 1:1 (ConA and PNA) or 1:2 (GNA) models. Frequency shifts (Δf) were recorded upon injection of 104 µg/ml of ConA, 104 µg/ml of GNA or 220 µg/ml of PNA. Molar concentrations were calculated considering molecular masses of 104 kDa for ConA, 52 kDa for GNA, and 110 kDa for PNA, and normalised $-\Delta f$ values were calculated by dividing $-\Delta f$ by the molecular mass of each lectin. Standard deviations of mean values are given.

4.3 Evaluation of the functionality of bacterial adhesins and microarray-assisted screening of anti-adhesive compounds

In the era of increasing antibiotic resistance, antivirulence has emerged as an alternative chemotherapeutic strategy, prompting the discovery of new bactericide or bacteriostatic molecules, as well as potential anti-adhesives (Escaich, 2008). Bacteria have evolved refined machineries aimed at actively invading target cells for replication and circumventing host immunity (Pizarro-Cerdá and Cossart, 2006; Kline *et al.*, 2009). Bacterial adhesion is one of the very first steps of the infective process and is often mediated by carbohydrate–protein interactions, as in the archetypal case of the *E.coli* mannose-specific adhesin, FimH (Mulvey, 2002).

As the interaction between a single sugar moiety and lectins is typically characterised by low intrinsic affinity, urinary tract infectious *E.coli* (strain UTI89) bears multiple copies of FimH, enabling the establishment of multivalent interactions with arrays of mannosylated uroplakin Ia present in the urothelium (Zhou *et al.*, 2001), thereby increasing binding avidity and guaranteeing a strong adherence. The cluster glycoside effect may result in an increase in avidity of several orders of magnitude (Lundquist and Toone, 2002), what is particularly important for lectins showing low binding affinities towards monovalent ligands (Reichardt *et al.*, 2013). Therefore, FimH multivalent ligands based on a variety of synthetic scaffolds (Chabre and Roy, 2013), as e.g. glycofullerenes, glycodendrimers, or nanodiamonds, have been synthesised and assessed as anti-adhesives using different approaches, such as SPR, ELISA, heamagglutination, or bladder cell assays (Heidecke and Lindhorst, 2007; Gouin *et al.*, 2009; Durka *et al.*, 2011; Barras *et al.*, 2013).

A novel microarray application in the field of preventive, anti-adhesive therapies is here presented. Microarrays were designed and developed for evaluation of the functionality of FimH in paraformaldehyde-fixed bacteria as well as in bacteria at different growth phases. A competitive binding assay was established, encompassing immobilisation of model multivalent ligands and the use of ConA for platform validation. Different mannose-bearing multivalent compounds were tested as inhibitors of ConA binding to the arrays, and SPR and ITC analyses of the interaction served to confirm the microarray-derived results. The validated microarray set-up was finally used for the

evaluation of glycofullerenes as inhibitors of bacterial adhesion to the array-printed ligands.

4.3.1 Establishment of an efficient microarray set-up for evaluation of the functionality of *E. coli* FimH adhesin

The applicability of the established bacteria microarrays for detecting functional FimH on the surface of *E. coli* UTI89 was first examined. Fixed and SYTO® 13-labelled bacteria were printed in both manual and robotic arrays and the binding of biotin-labelled *α*-heptyl-mannoside (MAF285, Fig. 3.13) at increasing concentrations was tested (up to 0.1 mM). No significant binding was detected, what could be in principle explained by a low affinity of FimH towards this monovalent ligand. Tested in parallel, moderate fluorescence signals were observed for the binding of this ligand to printed ConA (Fig. 4.40, *a*), although efficient inhibition (>75%) was not achieved for the higher ConA concentrations in the presence of *α*-Me-Man concentrations up to 0.2 M (Fig. 4.40, *b*). This observation suggested that non-carbohydrate-mediated interactions may occur upon NC-immobilisation. Similarly, bacteria printing could negatively have affected the functionality of FimH, conformational changes counting for its inability to bind the heptosyl-ligand (Rabe *et al.*, 2011).

Therefore, a different microarray approach was next designed (Fig. 4.40, *c*). Mannan from *S. cerevisae* along with *S. cerevisae* cells and RNase B, a glycoprotein bearing a single high-mannose glycosylation site, were manually printed on NC-coated slides. Indeed, the mannose-binding activity of *E. coli* adhesins was previously studied using yeast agglutination assays (Ofek and Beachey, 1978), also including determination of the inhibitory activity of mannose-based compounds (Firon *et al.*, 1983). Strong ConA

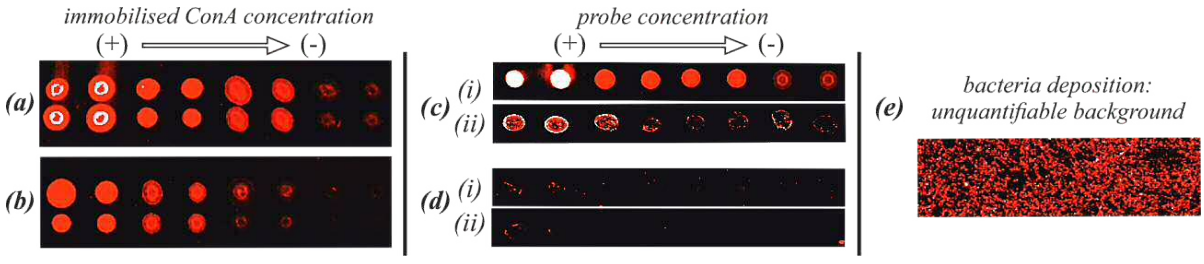


Fig. 4.40 Towards the establishment of an efficient microarray set-up for assessing FimH functionality. *(a)* Binding of 0.1 mM MAF285 to ConA, immobilised in quadruplicates at decreasing concentrations. *(b)* Competition experiment in the presence of 0.2 M *α*-Me-Man. *(c)* Binding of 0.4 μM biotinylated ConA to immobilised mannan *(i)*, or unstained *S. cerevisae* cells *(ii)* at decreasing concentrations (duplicates shown). *(d)* Respective competition experiments in the presence of 0.1 M *α*-Me-Man. *(e)* NC-slide incubation with SYTO 62®-stained UTI89 (OD_{600}=0.5). All experiments were carried out in HBS, pH 7.4.

binding to these arrays was observed, and it was inhibited in the presence of 0.1 M α-Me-Man (Fig. 4.40, *d*) confirming carbohydrate-dependent recognition. However, upon incubation with fixed SYTO® 62-stained UTI89, a thin film of bacteria was deposited over the whole NC membrane, even when continuous shaking was applied (Fig. 4.40, *e*). The consequent high background noise in the laser-scanned images precluded a quantitative or even qualitative assessment. Therefore this set-up was also discarded.

As alternative to NC-coated slides, the use of succinimide-activated hydrogel surfaces (H-slides) was then examined. Printing of neoglycoprotein ligands on hydrogel slides to enable bacterial capture had previously afforded quantifiable results, with low non-specific binding or surface adsorption (Adak *et al.*, 2013). Based on this precedent, mannan-poly-L-lysine (see section 3.5.1) and RNase B were immobilised on H-slides. Cy3®, which was successfully used for tracing probes upon arraying on NC-coated slides, was not useful in this case, since it is washed off during the deactivation and blocking steps (see 3.5.4.2.2). Thus, validation of the microarray set-up was carried out by examining the binding of ConA, as reliable binder of mannose-containing ligands.

The expected concentration-dependent trend was observed for ConA binding, evidencing that both mannan–poly-L-lysine and RNase B probes were efficiently immobilised in H-slides (Fig. 4.41, *a*). Moreover, the binding signals were stronger than those observed for NC-slides prepared using the same probe dilutions, pointing to higher immobilisation efficiency and/or favoured presentation in the hydrogel matrix

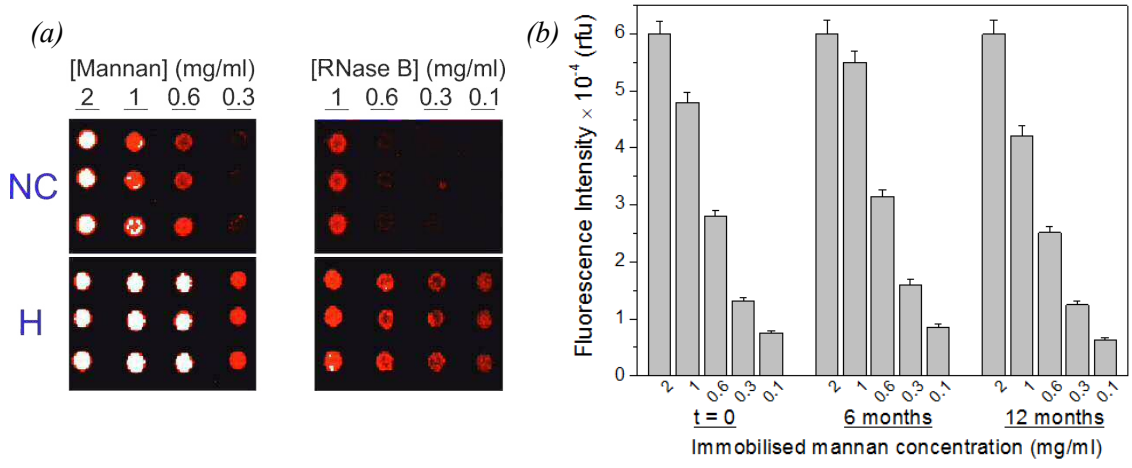


Fig. 4.41 ConA binding to mannan–poly-L-lysine and RNase B immobilized on NC and H microarray surfaces. (*a*)The probes were robotically printed in triplicates at decreasing concentrations and the binding of biotin-labelled ConA (0.4 µM) was monitored. White spots indicate detector saturation and correspond to mean rfu values >60,000. (*b*) Stability assessment of printed H-slides. Mannan–poly-L-lysine was printed as sextuplicates, and the binding of ConA (0.1 µM) was monitored immediately upon deactivation and blocking (t = 0), and after 6- or 12-month storage at -20 ºC. All experiments were carried out in HBS, pH 7.4.

for ConA binding over the NC arrays. Of note, comparable ConA binding to freshly prepared H-slides or after 6- or 12 month storage at -20 ºC was observed (Fig. 4.41, *b*), demonstrating that a set of simultaneously printed slides can be used for comparative testing of different lectin/bacteria targets for at least 1 year.

Having proved the applicability of the established H-arrays for lectin-binding assays, we next examined the binding of fixed SYTO® 62-stained UTI89 (harvested at the stationary growth phase) at different dilutions of OD_{600} up to 1. Bacteria deposits and film formation were not observed in this case. However, binding to the printed probes was neither observed (Fig. 4.42, *a*). A possible explanation was that FimH could be damaged by paraformaldehyde fixation, which might result in secondary structure "lock-in" or cross-linking aggregation (Hopwood, 1969; Mason and O'Leary, 1991). Down-regulation of FimH expression during the stationary phase of UTI89 could also account for the lack of binding.

Comparative binding experiments with live non-fixed bacteria in stationary and exponential growth phases revealed that both factors are decisive (Fig. 4.42). Thus, an aliquot of the bacterial culture was withdrawn during the logarithmic growth phase (OD_{600} = 0.3-0.5), and the rest was allowed to reach the stationary phase (OD_{600} above 1) before being harvested and stained. Live, non-fixed UTI89 in the exponential growth phase strongly bound to both RNase B and mannan–poly-L-lysine in a probe

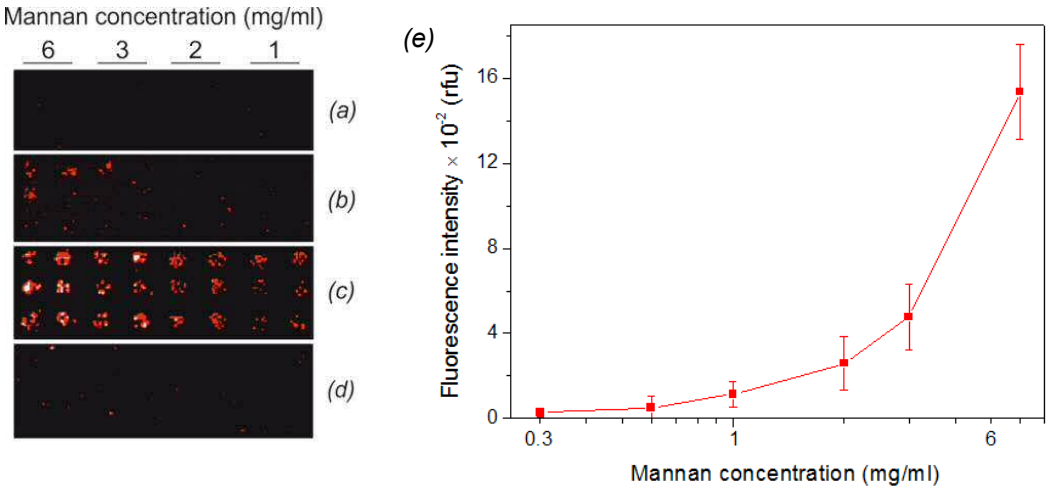


Fig. 4.42 Binding of UTI89 to mannan–poly-L-lysine printed on H-slides. The binding of SYTO 62®-stained UTI89 (OD_{600} = 1 in HBS, pH 7.4) to decreasing concentrations of mannan–poly-L-lysine, printed as sextuplicates, was tested. *(a)* Stationary phase fixed bacteria *(b)* Stationary phase non-fixed (live) bacteria *(c)* Exponential growth phase not-fixed (live) bacteria. *(d)* Inhibition of exponential growth phase bacteria binding in presence of 0.1 M *α*-Me-Man. *(e)* Mannan concentration dependence of the binding to the arrays of UTI89 in exponential growth phase. All experiments were carried out in HBS, pH 7.4.

concentration-dependent manner (Fig. 4.42, *c,e*), and the binding was successfully inhibited in the presence of 0.1 M *α*-Me-Man (Fig. 4.42, *d*), evidencing carbohydrate-mediated recognition. In striking contrast, live bacteria in the stationary growth phase only gave weak binding signals (Fig. 4.42, *b*), and paraformaldehyde-fixation completely abolished binding (Fig. 4.42, *a*).

Indeed, expression of type 1 fimbriae in *E. coli* is subject to phase-dependent control, and transcription of *fimA*, the gene encoding the fimbrial subunit protein is repressed as cells enter the stationary phase (Dove *et al.*, 1997). Thus, the established microarray set-up can be used to discriminate between fimbriate and afimbriate states by assessing FimH-mediated binding.

4.3.2 Microarray-assisted screening of anti-adhesive compounds

The efficient mannose-specific binding of live UTI89 to the mannan microarrays could also serve as platform for the evaluation of compounds as inhibitors of FimH-mediated bacterial adhesion. To explore the applicability of the approach, we again used ConA as model lectin, also allowing validation by other techniques of the microarray-derived results.

4.3.2.1 Testing of silica-core glyconanoparticles as inhibitors of ConA

Despite their widespread use for drug targeting (Zhou *et al.*, 2015), sensing or fluorescent detection (Hao *et al.*, 2016), the study of the interaction of lectins with silica nanoparticles (SNPs) covered with sugars started relatively recently (Wang *et al.*, 2011; Wang *et al.*, 2013; Kong *et al.*, 2015). In this Thesis, *α*-Me-Man- and *α*-Me-diMan-[Man*α*(1,3)Man-*α*-OMe] derivatised SNPs were synthesised as described in section 3.4, and evaluated for their binding capacity to ConA. Photo-initiated coupling of the underivatised sugars on PFPA-activated NPs, yielded GNPs carrying 14 and 4.3 μmol of *α*-Me-Man and *α*-Me-diMan-NPs, respectively, per g of particle, as quantified by the phenol/sulfuric acid method using bare NPs as blank. The quantity of immobilised *α*-Me-Man on the SNP-surface was comparable to that previously reported for Man- and Gal-conjugated SNPs (Kong *et al.*, 2015).

The activity of the GNPs as inhibitors of 0.4 μM ConA binding to mannan–poly-L-lysine and RNase Bin the Slide H set-up was first examined. However, no inhibition

was observed at any printed mannan/RNase B concentration, even when using GNPs concentrations up to 9 mg/ml. Thus, the binding avidity of ConA for immobilised, clustered ligands was noticeably superior to its affinity for the tested GNPs. Therefore, an alternative set-up was arranged, involving binding of unlabelled ConA to the mannan arrays followed by incubation with biotinylated heptyl mannoside (MAF285, see Fig. 3.13) in the absence or presence of GNPs. Because incubation with heptyl mannoside alone could result in inhibition of ConA binding to some extent, extensive control experiments were run in order to establish the most appropriate working concentrations

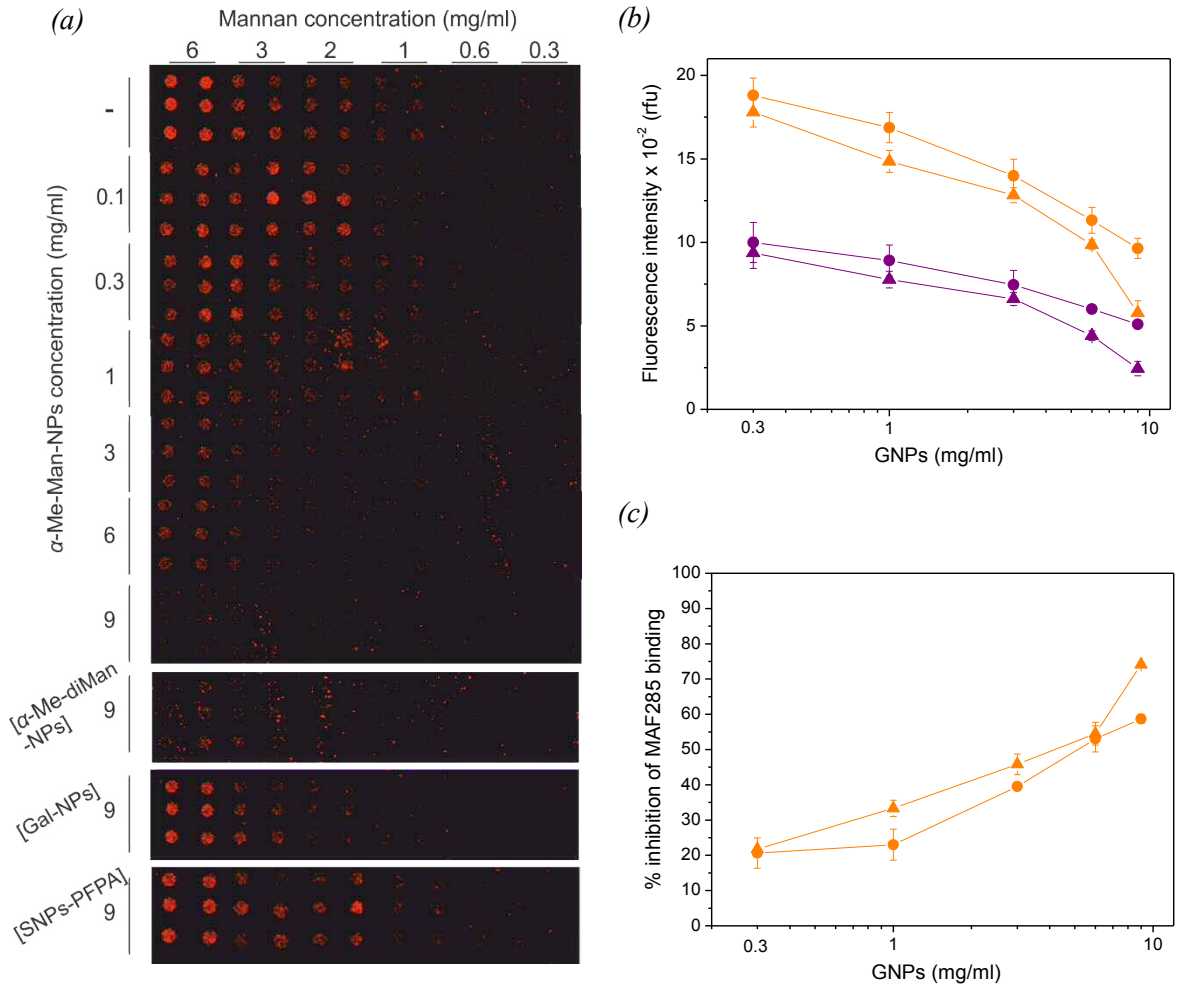


Fig. 4.43 Inhibition of heptyl mannoside-binding to ConA by GNPs. *(a)* ConA was anchored on mannan–poly-L-lysine printed as sextuplicates at different concentrations on hydrogel slides, and the binding of 0.1 mM MAF285 to ConA was next tested in the absence (–) and presence of α-Me-Man-, α-Me-diMan, or Gal-NPs and underivatised NPs (SNPs-PFPA). *(b)* Fluorescence intensity of MAF285 bound to ConA, anchored on 6 mg/ml (orange) or 3 mg/ml (purple) mannan–poly-L-lysine, in the presence of increasing concentrations of α-Me-Man- (triangles) or α-Me-diMan-NPs (circles). Data correspond to the mean fluorescence intensity of sextuplicates and the error bars to the standard deviation to the mean. *(c)* Inhibition percentage of MAF285 binding to ConA anchored on 6 mg/ml mannan–poly-L-lysine. Symbols are as in *(b)*. Data points represent the mean of at least two independent experiments and error bars correspond to standard error values. All experiments were performed in HBS, pH 7.4.

for both the lectin and MAF285. Using 0.8 µM ConA and 0.1 mM MAF285, fluorescence signals for MAF285 binding to ConA were in the linear range, enabling the evaluation of GNPs in competition experiments (Fig. 4.43). Both *α*-Me-Man- and *α*-Me-diMan-NPs were effective inhibitors of MAF285 binding to ConA, whereas neither Gal-NPs nor bare PFPA-derivatised NPs competed with MAF285 (Fig. 4.43, *a*), demonstrating the specificity of the inhibitory activity of the Man-derivatised NPs.

The inhibitory potential of *α*-Me-Man-NPs was slightly but consistently higher than that observed for *α*-Me-diMan-NPs (Fig. 4.43, *b*-*c*), what could be due to the smaller sugar content of the latter. Fifty per cent inhibition was observed around 3-6 mg/ml of GNP, which in terms of molar concentration of sugar corresponded to 4.2-8.4 × 10^{-5} M of *α*-Me-Man and 1.3-2.6 × 10^{-5} M of *α*-Me-diMan, respectively. In order to assess how these sugar concentrations compare with the binding affinity of ConA for the respective free sugars in solution, the association constants were determined by ITC.

Data obtained from the calorimetric titration of ConA with both saccharides could be fitted to a one-set-of-sites model, Fig. 4.44. The thermodynamic parameters derived from the analysis are given in Table 4.13. The stoichiometric values (*N*) suggested that only 80% of lectin subunits in the batch used for the analysis efficiently bound mannose. Both interactions were enthalpically driven, with an unfavourable entropic contribution, as usually observed for lectin–sugar interactions. The enthalpy change associated to the binding of the monosaccharide was lower than for the disaccharide, in line with previous reports (Williams *et al.*, 1992).

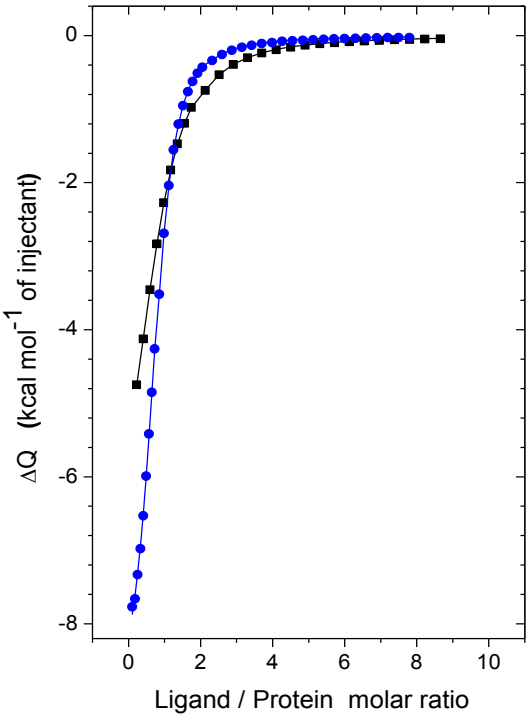


Fig. 4.44 Calorimetric titration of ConA with *α*-Me-Man (■) and *α*-Me-diMan (●) at 25 °C. ConA was loaded in the calorimeter cell at 270-300 µM in HBS, pH 7.4, and each ligand was injected in a stepwise fashion. Continuous lines correspond to data fit to a one-set-of-sites model.

More significantly, the binding affinity of ConA for *α*-Me-Man was 3.4-fold lower than for *α*-Me-diMan. Thus, the difference in affinity fully accounts for the comparable inhibitory activity of *α*-Me-Man- and *α*-Me-diMan-NPs on MAF285 binding to ConA in the microarray competition assays, despite the 3.2-fold difference in sugar content.

Table 4.13 Thermodynamic parameters for the binding of *α*-Me-Man and *α*-Me-diMan to ConA at 25 ºC

Compound	*N*	K_a (M^{-1})	ΔH ($cal \cdot mol^{-1}$)	ΔS ($cal \cdot mol^{-1} \cdot degree^{-1}$)
***α*-Me-Man**	0.772 ± 0.004	$(8.77 \pm 0.09) \times 10^3$	-7985 ± 55	-8.74
***α*-Me-diMan**	0.759 ± 0.002	$(3.02 \pm 0.04) \times 10^3$	-9370 ± 40	-10.9

On the other hand, assuming that the binding affinity of ConA for MAF285 is comparable to that measured for *α*-Me-Man, dissociation constants of 2.2-4.5 × 10^{-5} M and 0.7-1.4 x 10^{-5} M can be estimated for the *α*-Me-Man- and *α*-Me-diMan GNPs, respectively, which represent 2.5-5.2- and 2.4-4.7-fold enhancements in inhibitory activity of the GNPs over the free ligands. Still, these enhancements were not sufficient for achieving efficient inhibitory activity on the binding of ConA to immobilised ligands such as mannan–poly-L-lysine or RNase B. For detailed calculation of the dissociation constants here describe please see Annex V.

A random orientation of sugars when clicked on PFPA, the density or/and proximity of the silica core could be a possible explanation for the lack of a more significant clustering effect (Weber *et al.*, 2014) that could favour ConA recognition. Indeed, appropriate presentations should result in affinity enhancements of more than one order of magnitude in the case of monovalent ligands (Weber *et al.*, 2014). Thus, future work should address the optimisation of the photo-initiated coupling mechanism for functionalisation of PFPA-SNPs. At any rate, since the silica-based GNPs did not prove to be suitable candidates as efficient inhibitors of FimH-mediated bacterial adhesion, we next tested different glycofullerenes as alternative multivalent compounds with well-characterised sugar-derivatisation, again using ConA for initial evaluation.

4.3.2.2 Testing of glycofullerenes as inhibitors of ConA

Four fullerene hexakis-adducts, synthesised at the laboratory of Prof. Stéphane Vincent (University of Namur/FUNDP, Belgium), were evaluated (Fig. 4.45), three of them

carrying 12 peripheral *α*-Man residues grafted onto the C_{60} core by different spacers (compounds 1-3), and the fourth one bearing 12 peripheral Gal moieties (compound 4), to be used as negative control. Glycofullerenes were obtained by a "click" reaction between monovalent azido-sugars and the trimethylsilyl-protected polyalkyne fullerene scaffold, Fig. 4.45 (Iehl and Nierengarten, 2009; Nierengarten and Nierengarten, 2014), and their structure was confirmed by ^{1}H and ^{13}C NMR, infrared and mass spectroscopy (Abellán Flos, 2015).

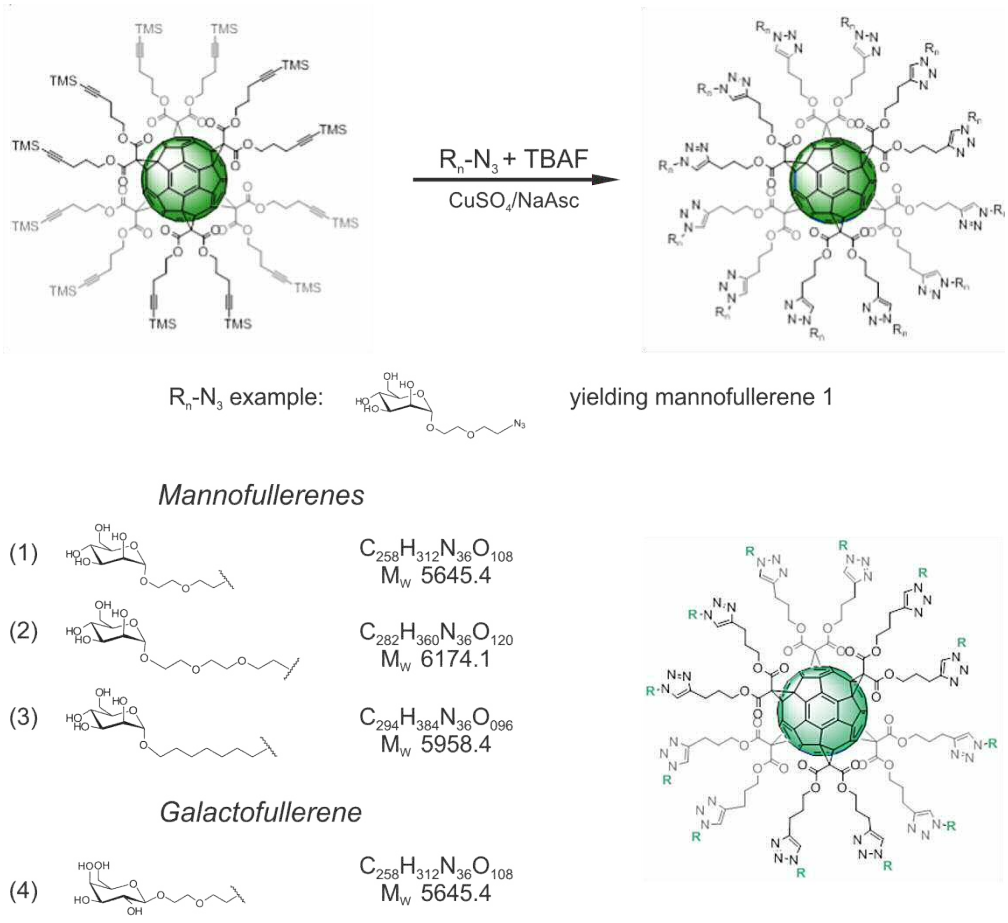


Fig. 4.45 Synthesis and illustration of the tested fullerene hexakis-adducts, bearing 12 peripheral hexose moieties. "Clicking" of monovalent azido-sugars (R_n-N_3) on the trimethylsilyl (TMS)-protected polyalkyne fullerene scaffold was assisted by tetrabutylammonium fluoride (TBAF) in presence of $CuSO_4$/NaAsc (0.1/0.3 *eq*) in THF/DMSO/H2O (3:3:1), 72 h at room temperature.

The heptosyl-mannose residue (compound 3), acknowledged to be an excellent ligand for FimH, was included to enable comparisons with existing literature (Bouckaert *et al.*, 2005; Compain *et al.*, 2010; Nierengarten *et al.*, 2010; Durka *et al.*, 2011), although its high hydrophobicity and poor water-solubility was reported to be disadvantageous. Therefore, flexible polyethyleneglycol (PEG) spacers of distinct length were employed for glycofullerenes 1, 2, and 4.

4.3.2.2.1 Microarray assay

Glycofullerenes 1, 2, and 4 were first tested as inhibitors of the binding of biotinylated ConA (0.4 μM) to mannan–poly-L-lysine and RNase B printed on H-slides, using fullerene concentrations up to 0.1 mM. Fullerene 3 was not tested in this set-up due to its poor solubility and limited availability. As the derived fluorescence signals for ConA-binding reached plateau for mannan concentrations above 0.3 mg/ml (also see Fig. 4.2, *a*), but were in linear range for all four tested concentrations for RNase B, the latter was selected for the evaluation of the inhibitory potential of the glycofullerenes.

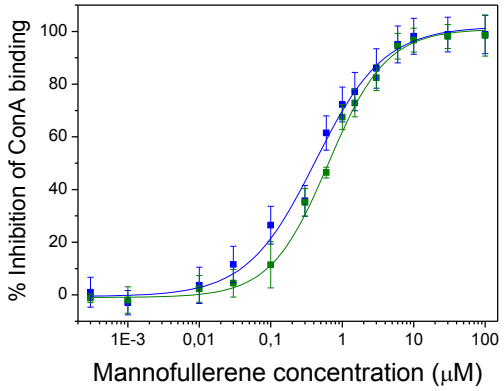


Fig. 4.46 Inhibition of ConA binding to RNase B by mannofullerenes 1 and 2. The binding of biotinylated ConA (0.4 μM) to RNase B printed at 1 mg/ml on H-slides was tested in the absence and presence of increasing concentrations of mannofullerenes 1 (green) or 2 (blue) in HBS, pH 7.4. For calculation of IC_{50} values, the results were fitted to a sigmoidal curve.

A dose-dependent inhibition by mannofullerenes 1 and 2 was observed (Fig. 4.46), whereas the galactofullerene had no effect, in agreement with ConA's carbohydrate-binding specificity. Sigmoidal fitting yielded inhibition curves from which IC_{50} values in the submicromolar range were estimated (Fig. 4.46 and Table 4.14). Interestingly, the IC_{50} value obtained for mannofullerene 2 was slightly below that obtained for mannofullerene 1, suggesting that the longer and flexible PEG linker may provide

Table 4.14 Inhibitory capacity of the compounds tested on ConA binding to RNase B

Compound	IC_{50} (M)	Relative inhibitory potential	
		Relative potency[a]	Normalised potency[b]
Mannofullerene 1	$(6.41 \pm 0.07) \times 10^{-7}$	713	59.4
Mannofullerene 2	$(4.20 \pm 0.09) \times 10^{-7}$	1088	90.7
***a*-Me-Man**	$(4.6 \pm 0.5) \times 10^{-4}$	1	1

[a] Relative potency taking the IC_{50} value obtained for α-Me-Man as unit. [b] Relative potency normalized per Man moiety.

additional spatial freedom and/or less steric hindrance to the attached mannose moieties, thereby facilitating the binding to ConA.

The inhibitory capacity of both mannofullerenes was compared to that of α-Me-Man, which was included as reference in all series of measurements. Of note, the IC_{50} value obtained for the monovalent hapten (457 μM) was similar to those derived from enzyme-linked lectin assays in plate format (530-870 μM), also using mannan as the immobilised ligand (Roy *et al.*, 1998; Köhn *et al.*, 2004). Thus, a strong multivalency effect was observed for the mannofullerenes tested, with an increase of up to 90-fold in inhibitory potency when normalised per Man moiety (Table 4.14). In order to confirm direct binding of ConA to mannofullerenes 1 and 2, we next performed SPR analyses of ConA-fullerene interactions.

4.3.2.2.2 Surface plasmon resonance analyses

SPR has been widely employed to investigate the different factors contributing to the binding enhancement arising from the glycocluster effect, using a variety of experimental set-ups (Mann *et al.*, 1998; Hernáiz *et al.*, 2002; Munoz *et al.*, 2013; Bayón *et al.*, 2016). In this work, ConA was immobilised at acidic pH on Biacore® sensor CM5 chips, as described in section 3.6.2, yielding an approximate response of 2000 RU. The functionality of immobilised ConA was confirmed by injecting 4.3 μM RNase B) over the chip, what resulted in a 100 RU increase. Mannofullerenes 1 and 2, as well as the galactofullerene as negative control, were then systematically tested under identical conditions in at least two independent experiments using different chips. Mild regeneration conditions were optimised to yield at least 95% of surface regeneration with a single injection.

Mannofullerenes 1 and 2 gave positive and concentration dependent signals (Fig. 4.47, left), the interaction apparently following pseudo-first order kinetics, whereas injection of galactofullerene resulted in no response (Fig. 4.47, right), proving the absence of non-specific interactions. Maximal response data were plotted versus the fullerene concentration (Fig. 4.48) and, the dissociation constants (K_D) were calculated in equilibrium following a 1:1 interaction model (see section 3.6.2), yielding satisfactory Chi^2 values (Table 4.15).

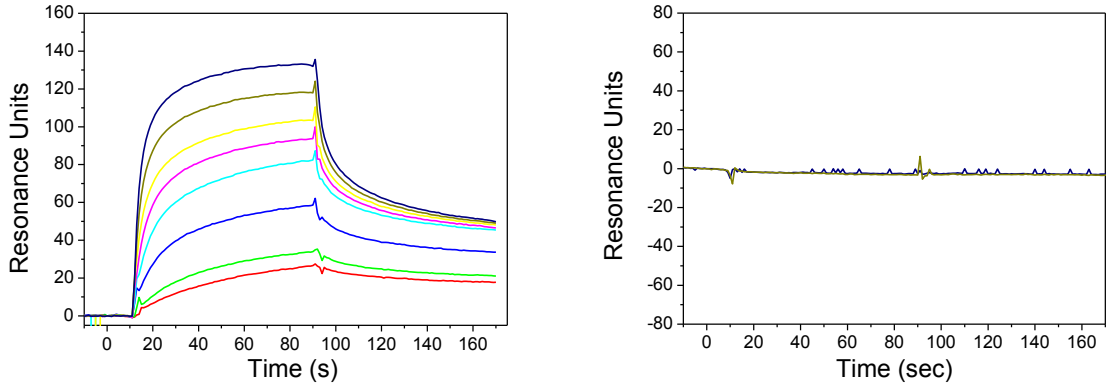


Fig. 4.47 Representative sensograms for the binding of mannofullerene 1 and galactofullerene to ConA immobilized on CM5 chips. Left panel: mannofullerene 1was injected at 0.25 (red), 0.5 (green), 1 (blue), 2 (cyan), 3 (magenta), 5 (yellow) 7.5 (dark yellow), and 10 μM (navy) in HBS, pH 7.4, followed by 100 s of buffer flow, corresponding to the dissociation phase, and a single 20-s injection of 10 mM glycine-HCl, pH 3.0, achieve at least 95% regeneration of the surface. Right panel: for galactofullerene, only the two highest concentrations, 7.5 (dark yellow) and 10 μM (navy) were tested, resulting in no response. Performed at 22 °C.

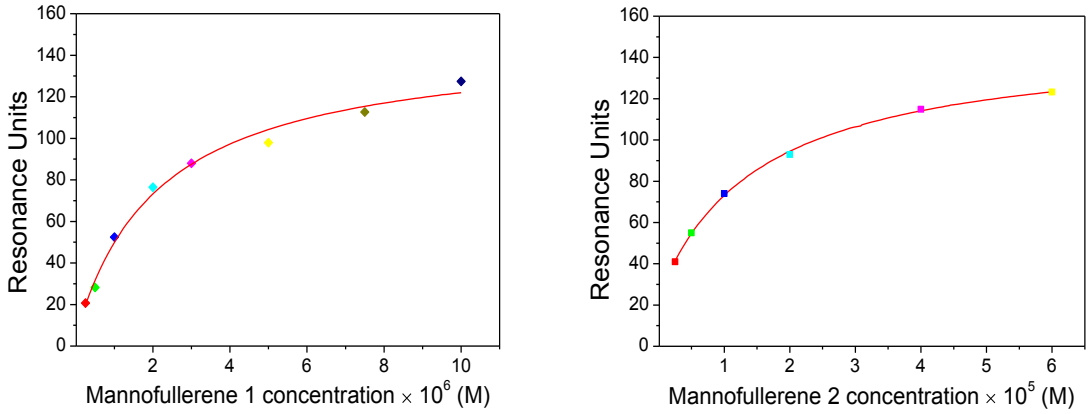


Fig. 4.48 Maximum resonance shift detected for the binding of mannofullerenes 1 and 2 to ConA immobilised on CM5 chips. Mannofullerenes 1(left panel) and 2 (right panel) were injected at different concentrations, as described in the legend to Fig. 4.47, and the maximum response was plotted as a function of fullerene concentration.

Table 4.15 SPR-derived parameters of mannofullerene binding to immobilised ConA

Mannofullerene	K_D (M)	R_{max} (RU)	$Chi^2(RU^2)$
1	$(2.1 \pm 0.4) \times 10^{-6}$	132 ± 5	6.91
2	$(1.5 \pm 0.1) \times 10^{-5}$	125 ± 2	1.38

Data processing and analysis were carried out with Biacore X-100 evaluation software applying the one-set-of-sites model (1:1).

The obtained K_Ds were within the sensitivity limits reported for the Biacore® instrument used in this study, i.e. 10^{-3}-10^{-6} M (Hall *et al.*, 1996). Unexpectedly, the K_D for mannofullerene 1 was one order of magnitude lower than for mannofullerene 2, a trend opposite to that observed in the microarray set-up, in which both ConA and the mannofullerene are in solution. A similar trend was witnessed before by an SPR competition assay applied in the evaluation of the ConA-inhibitory potential of mannose-derivatised gold NPs bearing linkers of distinct length (Lin *et al.*, 2003).

At any rate the longer PEG spacer of mannofullerene 2 apparently disfavours the establishment under flow of multivalent contacts with ConA immobilized on the SPR chip, whereas it seems to facilitate the binding when the lectin is also in solution (microarray assay). To evaluate this hypothesis and in order to determine the thermodynamic parameters of the interaction, an ITC analysis was carried out.

4.3.2.2.3 Isothermal titration calorimetry analysis of ConA-mannofullerene binding

ConA (at 79-90 μM monomer concentration) was titrated with 0.2 mM solutions of glycofullerenes at the temperature used for the microarray and SPR analyses, *i.e.* 22 ºC. Titration with galactofullerene yielded small endothermic ITC signals (Fig. 4.49) similar to those obtained for compound dilution, as expected for the absence of interaction.

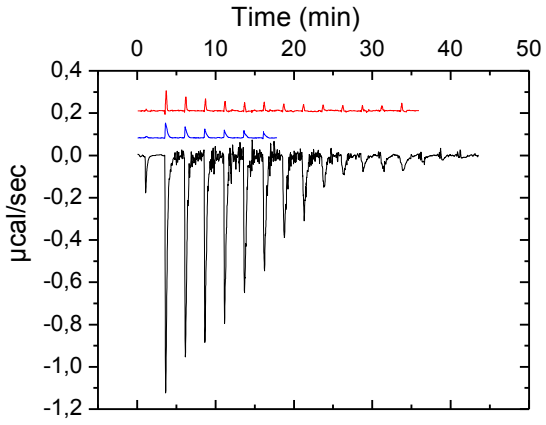


Fig. 4.49 Calorimetric titration of ConA with glycofullerenes.
Representative raw calorimetric data obtained upon 18 automatic injections of mannofullerene 1 (black), into the cell containing 90 μM ConA. The blue peaks correspond to the heat of dilution of the mannofullerene when injected into the cell filled with buffer. Red peaks correspond to raw calorimetric data obtained for the titration of ConA with galactofullerene. All the titrations were carried out in HBS, pH 7.4, at 22ºC.

In contrast, significant exothermic signals were obtained upon titration of ConA with mannofullerenes 1 and 2 (Fig. 4.49 and 4.50). Titration curves of several independent experiments are shown in Fig. 4.50. For calculation of the binding parameters, the "ligand-in-cell" analysis option was used, enabling the determination of the number of binding sites for ConA per mannofullerene molecule. Mean thermodynamic parameters derived from at least three independent experiments are compiled in Table 4.16. Interestingly, the number of binding sites for ConA per fullerene varied between 6-8 for mannofullerene 1 and 6-9 for mannofullerene 2, depending on the experiment, indicating that simultaneous binding of ConA to the 12 mannose residues per molecule was not possible, most likely due to steric hindrance effects. It has been reported for the ConA tetramer that it presents two sites on each face, the distance between them being 6.5 nm (Derewenda *et al.*, 1989). Worth mentioning, at pH 7.4 ConA forms bulky tetramers (Senear and Teller, 1981) that might obstruct access to neighbour mannose residues. In the same context, tetramerisation also qualifies ConA as multivalent

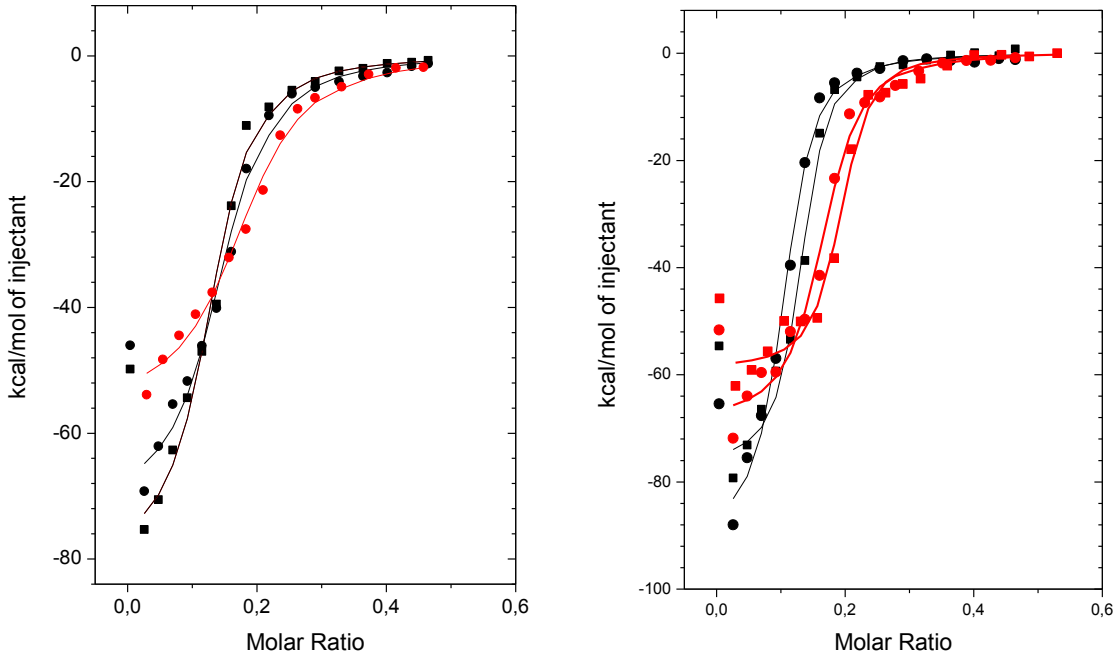


Fig. 4.50 Analysis of data derived from calorimetric titration of ConA with mannofullerenes 1 and 2. Two different preparations of ConA at 79 (red) and 90 μM (black) were used for the titration with mannofullerenes 1 (right panel) and 2 (left panel). Circles and squares correspond to experimental data obtained for two independent experiments, and the solid lines to the fit of the data using a one-set-of sites model and the "ligand-in-cell" analysis option. The mean of functional monomer concentration *N*, calculated for the used batch of ConA , was applied as input for ligand concentration in the fitting of the dodecamannosylated fullerene titrations. All experiments were carried out in HBS, pH 7.4, at 22 ºC.

receptor, thus enabling the formation of cross-linked complexes. Indeed, samples recovered from the calorimeter cell upon completion of the experiments were cloudy, proving extensive cross-linking at saturation.

The enthalpy changes (ΔH) associated to the binding of both mannofullerenes were identical, but the entropic penalty (ΔS) was smaller for mannofullerene 2, resulting in a 2.9-fold higher association constant, in line with the tendency observed in the microarray competition experiment. This observation concerning the length of linker was consistent with the reported by others (Chien *et al.*, 2008; Wang *et al.*, 2010) for comparable systems and experimental conditions (in-solution evaluation).

Table 4.16 Thermodynamic parameters for the binding of ConA to mannofullerenes 1 and 2 and to *α*-Me-Man and the monovalent Man-derivative bearing the spacer of mannofullerene 1 (compound 1a)

Compound	N^a	K_a (M^{-1})	ΔH (cal/mol)	ΔS (cal/mol/degree)
Mannofullerene 1	6.8 ± 0.2	$(1.5 \pm 0.3) \times 10^5$	$-(1.01 \pm 0.03) \times 10^4$	-10.63
Mannofullerene 2	7.4 ± 0.2	$(4.3 \pm 0.6) \times 10^5$	$-(1.01 \pm 0.03) \times 10^4$	-8.88
***α*-Me-Man**	0.862 ± 0.002	$(9.48 \pm 0.05) \times 10^3$	$-(7.74 \pm 0.03) \times 10^3$	-7.99
1a	1.45 ± 0.01	$(7.0 \pm 0.1) \times 10^3$	$-(5.21 \pm 0.04) \times 10^3$	-0.05
1a[b]	1	$(7.0 \pm 0.1) \times 10^3$	$-(7.55 \pm 0.04) \times 10^3$	-7.98

[a] For mannofullerenes 1 and 2, *N* represents the number of mannose residues occupied by ConA per dodecamannosylated fullerene. For *α*-Me-Man, *N* corresponds to the relative number of functional ConA monomers calculated for that specific batch of lectin. For compound 1a, *N* corresponds to the number of molecules calculated to be recognised per ConA site. Except for the second set of values given for 1a, K_a, ΔH, and ΔS (cal/mol/degree) are calculated per mole of sugar. [b] Recalculation of thermodynamic parameters per mole of ConA monomer. Data for mannofullerenes 1 and 2 correspond to the mean of at least three individual experiments.

In order to assess the effect of glycofullerene's multivalent presentation of Man residues on the binding affinity towards ConA, titrations were also performed using *α*-Me-Man and the Man-derivative serving as precursor of mannofullerene 1 (*i.e.*, the azido-sugar containing the PEG spacer of this mannofullerene, designated as compound 1a) (Fig. 4.51). The thermodynamic parameters obtained for the binding of *α*-Me-Man to the specific ConA preparation used in this case (Table 4.16) were very similar to those reported in Table 4.I, thus validating the results of the two sets of ITC experiments. The K_a was found to be 16-45-fold lower than that measured for the mannofullerenes, in line with the relative potency trend observed in the microarray competition assays.

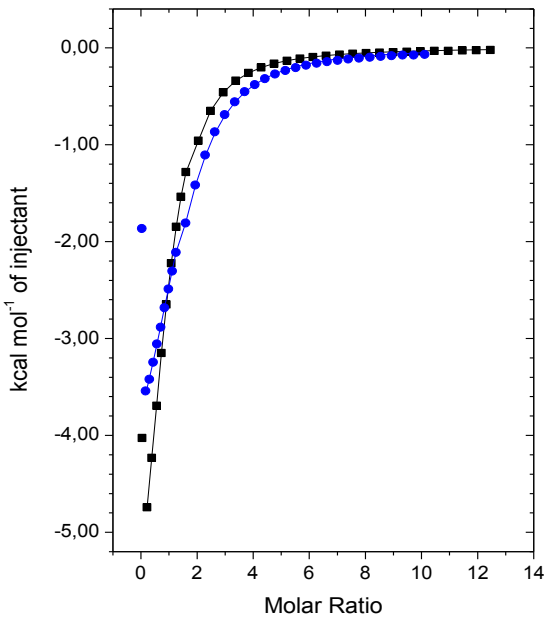


Fig. 4.51 Calorimetric titration of ConA with *α*-Me-Man and compound 1a. Con A (239 μM) was titrated with 15 mM *α*-Me-Man (■) or compound 1a (●) in HBS, pH 7.4 at 22 ºC. Continuous lines correspond to the fit to one-set-of-sites model.

The K_a value obtained for compound 1a was comparable to that obtained for *α*-Me-Man (Table 4.16). However, the binding stoichiometry approached 1.5 molecules of 1a per ConA subunit, resulting in lower enthalpy and entropy changes per mole of sugar. This behaviour could be attributed to self-association or formation of an aggregate or other self-assembled structure, as recently brought up elsewhere (Kong *et al.*, 2016). Indeed, the capacity of compounds bearing oligo(ethyleneglycol) linkers to form self-assembled monolayers has been described (Harder *et al.*, 1998). On these terms, the minimal (if existent) entropic cost of the ConA interaction with compound (1a) would argue for the formation of a higher-order assembly. Recalculation of ΔH and ΔS per mole of ConA monomer yielded values very similar to those obtained for *α*-Me-Man, indicating that the presence of the PEG spacer does not affect recognition of the sugar moiety. Moreover, the results evidenced that grafting of 1a onto the fullerene core resulted in an approximate 20-fold enhancement in the binding affinity.

Overall, ITC data fully supported the binding trends observed in the microarray competition experiments, with both mannofullerenes exhibiting a noticeably higher inhibitory activity compared to the free monosaccharide. Thus, having validated the microarray set-up for evaluation of the glycofullerenes as inhibitors of ConA, we were in a position to examine their inhibitory potential on *E. coli* binding to mannan microarrays.

4.3.2.3 Testing of glycofullerenes as *E. coli* anti-adhesives

The binding of live SYTO®62-stained UTI89, in exponential growth phase, to mannan printed on H-slides was examined in the absence and presence of increasing concentrations of glycofullerenes, this time including mannofullerenes 1, 2, and 3, together with the galactofullerene as control for specificity.

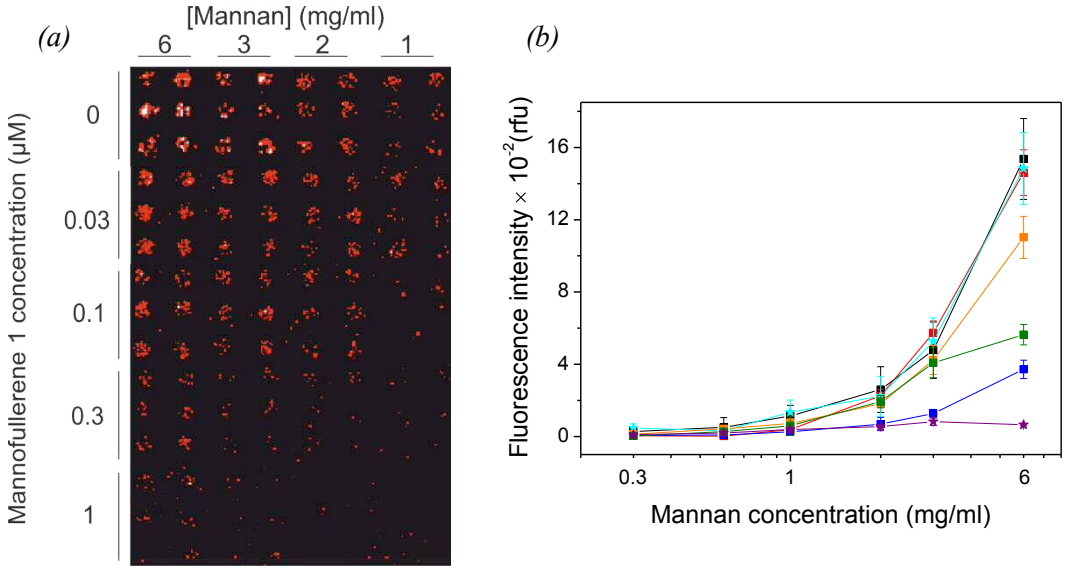


Fig. 4.52 Inhibition of the binding of FimH-expressing UTI89 by mannofullerenes. *(a)* The binding of live SYTO®62-stained bacteria in exponential growth phase to different concentrations of mannan–poly-L-lysine, printed as sextuplicates, was tested in the absence and presence of increasing concentrations of mannofullerene 1. *(b)* Fluorescence intensity of bacteria bound to mannan, printed at increasing concentrations, in the absence (■) or presence of 0.3 µM of mannofullerenes 1 (■), 2 (■), and 3(■), and of galactofullerene (■). *α*-Me-Man at 0.3 µM (★) or 10 mM (★) was used as reference.

As illustrated in Fig. 4.52 *a,* for mannofullerene 1, the binding was efficiently inhibited by mannofullerenes at concentrations as low as 0.3 µM. In comparison, the presence of 0.3 µM *α*-Me-Man produced no meaningful inhibition (Fig. 4.52 *b*), evidencing the increase in inhibitory activity resulting from the multivalent presentation of the sugar on the fullerene scaffold. In contrast, the galactofullerene did not show any inhibitory effect, in agreement with the specificity of FimH towards mannosides (Bouckaert *et al.*, 2005). Moreover, this result evidenced that neither the fullerene core nor the PEG spacer (the same as for mannolullerene 1), are involved in non-specific interactions that could affect FimH functionality. Thus, the inhibitory activity of the mannofullerenes can be confidently attributed to specific carbohydrate-mediated interactions.

Inhibition curves were built by semi-logarithmic plotting of the percentage of inhibition of UTI89 binding to mannan printed at 6 mg/ml versus the concentration of

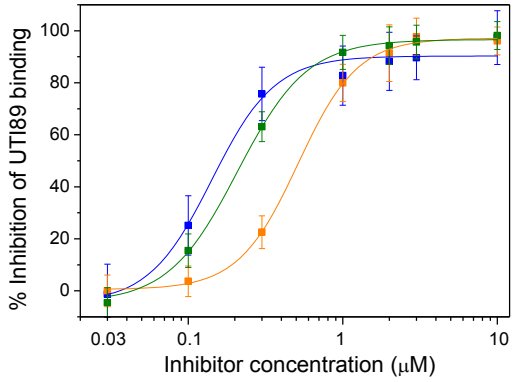


Fig. 4.53 Inhibitory activity of mannofullerenes on UTI89 FimH-mediated binding to immobilised mannan–poly-L-lysine. Bacteria binding to the highest mannan concentration (6 mg/ml) was tested in the presence of increasing concentrations of mannofullerene 1 (■), 2 (■), and 3(■). Data shown are the mean of at least four different experiments and error bars indicate its standard deviation. Solid lines correspond to a sigmoidal fit to experimental data.

mannofullerene, and sigmoidal fit to the data (Fig. 4.53). IC_{50} values were estimated from these curves (Table 4.17). The three mannofullerenes exhibited a clearly enhanced inhibitory activity compared to *α*-Me-Man, with 5 to 18-fold higher relative potencies per mannose residue. Mannofullerene 3 was 2.6-3.6-fold less potent inhibitor than mannofullerenes 1 and 2, which contain flexible PEG spacers. Moreover, mannofullerene 2 was slightly more efficient than mannofullerene 1, which bears a shorter spacer, following the same trend than in the case of ConA. Thus, the results indicate that mannose grafting to the fullerene core through flexible and appropriately long spacers facilitates binding to FimH, a lectin that presents a deep binding pocket (Bouckaert *et al.*, 2005; Durka *et al.*, 2011).

Table 4.17 Inhibitory capacity of the compounds tested for UTI89 binding to immobilised mannan

Compound	IC_{50} (M)	Relative inhibitory potential	
		Relative potency[a]	Normalised potency[b]
Mannofullerene 1	$(1.99 \pm 0.06) \times 10^{-7}$	154	13
Mannofullerene 2	$(1.42 \pm 0.02) \times 10^{-7}$	215	18
Mannofullerene 3	$(5.2 \pm 0.2) \times 10^{-7}$	59	5
***α*-Me-Man**	$(3.1 \pm 0.2) \times 10^{-5}$	1	1

[a] Relative potency calculated taking the IC_{50} value for α-Me-Man as unit. [b] The normalised potency corresponds to the relative potency divided by the number of sugar residues per fullerene molecule (12).

DISCUSSION

5.1 Pathogen-derived glycosylation patterns: typification and prospects for the establishment of virulence correlations

Bacteria-based microarrays have proved to be an efficient high-throughput tool for screening glycosignatures of NTHi, a highly variable pathogen, lacking capsular polysaccharides and large O-antigen-containing LPSs, thereby expanding the applicability of an approach previously developed in our laboratory (Campanero-Rhodes *et al.*, 2015). Besides the semi-quantitative data derived from the microarray binding analysis as for binding avidity, detailed information on the affinity and kinetics of lectin binding to entire bacteria cells could be obtained using the tailored QCM bacteria chips and real-time binding assays with label-free lectins (Kalograiaki *et al.*, 2016). Evidently, the two methods benefitted from each other, the microarray screening pointing to relevant candidates for bacterial capture and kinetic analysis, and the QCM providing real-time insights into the interactions occurring at the native bacterial surface, thus closely reflecting the real lectin-bacteria interplay.

Since bacterial virulence often correlates with the carbohydrate structures displayed and their recognition, or lack of recognition, by lectins of the innate immune system (Sahly *et al.*, 2008), the microarray approach proved to be a useful tool for exploring these features. In the case of NTHi, the extensive inter-strain heterogeneity in glycosylation patterns is likely to be strategic for its commensal behaviour, the comparison of concrete glycosignatures and their recognition by endogenous lectins aiming to aid in the establishment of functional and clinical correlations. In the microarrays the lectin-binding fingerprint for the tested isolates and mutants revealed the presence of specific carbohydrate structures to be potentially recognised by lectins of the innate immune system with the appropriate carbohydrate-binding specificity. As significant example, recognition of NTHi clinical isolates by WGA followed a similar trend as Siglec-14 (section 4.1.8) both reported to recognise Neu5Ac. A comparable correlation was observed for the Lac-specific RCA and the hGal-8 with the illustrative exception of NTHi1566. This event claims that the final presentation of glycans on the pathogen surface may count for the fine-tuning of endogenous lectin action as PRRs in the context of epitope clustering or, their protein or lipid scaffolding on the tissue or pathogen surface (Dam and Brewer, 2010).

In this context, the classification of galectins, long-established "self"-element binders, as PRRs may require additional considerations. Most metazoans are endowed with a complex galectin repertoire, members exhibiting multiple isoforms and subtle variations in carbohydrate specificity (Vasta *et al.*, 2012), also being well accepted that their direct binding to glycoconjugates on the surface of bacteria results in either positive or negative regulation of the innate and adaptive immunity (Vasta, 2009). Then again the molecular basis of their binding to distinct self and non-self glycans via the same CRD remains to be elucidated. In this regard, galectins with tandemly arrayed CRDs stand out as intriguing entities both in their binding properties and functional aspects, the most striking example being the recognition and killing of *E. coli* O86 displaying B-blood group oligosaccharides (BGB+) by the hGal-4 and -8 (Stowell *et al.*, 2010). The results obtained in the context of this Dissertation revealed that the C-ter. domain mediates recognition of NTHi375 by hGal-4, and that bacterial viability is not affected even when integrity of the full length galectin is guaranteed (section 4.1.10.3). Study of the role of the full-length hGal-4 in *Pseudomonas aeruginosa* (PAO1) susceptibility to antimicrobial peptides revealed increased cooperativity, despite its failure of direct killing (Buzamet, 2012). It is here suggested that the redox properties of the intracellular and extracellular environments may modulate the activity of lectin–ligand interactions, and the functional outcome. Most galectins are active in the intracellular, reducing environment, but once secreted they are susceptible to inactivation by oxidation of their free Cys residues present in the CRD (Stowell *et al.*, 2009), compromising the oligomerisation of the galectin subunits. Intriguingly, we observed a striking increase in hGal-4 recognition and NTHi375 agglutination in absence of reducing agent, possibly underscoring the formation of higher-order structures (section 4.1.10.4).

Apart from the mechanistic viewpoint, ligand assignment for lectin binding to the bacterial surface is of particular interest also for establishing virulence correlations. The results exposed in section 4.1 support that the dissection of specific epitopes for lectins is greatly facilitated when isogenic mutants or relevant strains are available. A parallel comparison of lectin-binding profiles of wild type *K. pneumoniae* strain Kp52145 and a panel of isogenic mutants lacking different surface epitopes (i.e. the capsule, LPS O-polysaccharide chain, or the major outer membrane protein OmpA), pointed to the O-antigen as the primary epitope recognised by the Gal-specific lectins tested (Campanero-Rhodes *et al.*, 2015). Consistently, hGal-3 is projected to recognise the

repetitive motif Gal*β*(1,3)Gal– on the O-antigen of purified Kp52145-LPS in a carbohydrate-specific manner, binding being completely abolished in presence of 10 mM hapten inhibitor, Gal*β*(1,3)Gal–OMe, in LPS-based microarrays (not shown). The *K. pneumoniae* LPS was previously shown to be also a ligand for rat Gal-3 (Mey *et al.*, 1996), sharing high homology with the human protein. Accordingly, for NTHi375 the LOS ramification off Hep III behaved as ligand for hGal-8N and the full length-protein, a stronger binding to Δ*lgtF* over Δ*ompP5* LOS being observed and truncation of the LOS extension at Hep III in Δ*lpsA* dramatically reducing recognition (Fig. 4.32).

Similarly, the results derived from lectin screening of the NTHi375 mutant library bearing sequentially truncated LOS, pointed to its being a docking site for Gal- and manno-Hep-binding lectins (section 4.1.3). Consistently, direct VAA and ConA binding was confirmed upon isolation of the LOS molecule and development of LOS microarrays (section 4.1.5.2). In the illustrative case of ConA and SP-D presenting similar preference as for manno-Hep (Jaipuri *et al.*, 2008; Wang *et al.*, 2008), strong binding to deep LOS-truncated mutants (Δ*lgtF*, Δ*lpsA* and double mutant) and the isolated molecules, indicated a protected target for both lectins to lay in the LOS inner core structure (section 4.1.9). What is more, our results provided the first experimental evidence for direct binding of SP-D to the surface of NTHi and also advocated for the proximal manno-Hep moiety (Hep I) being preferentially recognised by the collectin. The elucidation of the crystal structure of the complex of SP-D with the LOS derived from *Haemophilus influenzae* Eagan (Clark *et al.*, 2016) was fully consistent with these observations, suggesting the ramifications extension and complexity as key to efficiently disguise the vulnerable inner-core components of the LOS pattern.

On the other hand, besides its variable LOS, NTHi also presents a collection of adhesins and glycoproteins that constitute virulence factors of equal importance, and are associated to bacterial attachment and consequent invasiveness (Rao *et al.*, 1999). In particular, the adhesins HMW (High Molecular Weight glycoproteins 1/2) are present in the NTHi375 surface and glycosylated by HMWC following an alternative cytoplasmic pathway by adding mono- or di-Hex (mostly Glc- and Gal-) to certain Asn residues (Gross *et al.*, 2008; Grass *et al.*, 2010; Gawthorne *et al.*, 2014; Naegeli *et al.*, 2014). Remarkably, a recent experimental study highlighted how the transfer of the *hmw1* gene from 86028-NP (invasive isolate) to Rd KW20 (considerably less adherent) can alter its phenotype and increase its colonisation potential (Mell *et al.*, 2016). In our study we

used Rd strain and a mutant expressing HMW1 of NTHi12 and our lectin-screening results claimed that $HMW1_{12}$ does not constitute a ligand for RCA (section 4.1.6). Even though it is suspected that HMW1 does not present Gal in a favourable fashion for RCA recognition, such possibility should not be generally discarded. Indeed, HMW genes throughout the species share homology, but sequence alignments with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) revealed nine aminoacid substitutions for $HMW1_{375}$ with respect to $HMW1_{12}$ knocking out reported glycosylation sites (Domenech *et al.*, 2016), thus suggesting relevant heterogeneity among strains (see Annex VI). The binding of ConA was also significantly enhanced upon expression of the $HMW1_{12}$, a remark particularly expected since the presence of Glc and Glc-Glc units decorating the protein was formerly established (Grass *et al.*, 2003). Although previous evidence of direct ConA binding to HMW is lacking, it has been suggested that differences observed in ConA-labelling of biofilms derived from NTHi strains carrying *hmw* genes, including NTHi375 (Domenech *et al.*, 2016) could be furtherly attributed to variations in the expression of genes, known to be modulated by phase variation (Dawid *et al.*, 1999).

Membrane glycoproteins like HMW are actually hard to purify, thus limiting the study of molecular recognition reactions in solution. Entire bacteria microarrays proved out to be a powerful tool for the study of such systems, revealing that RCA and VAA despite being highly homologous and terminal Gal- specific, recognised different targets on the bacterial surface. Since the isolation of LOS was possible, STD-NMR experiments have served to furtherly define the LOS(s) epitopes recognised by these two lectins at the atomic level, supporting our previous findings. On the LOS derived from NTHi375Δ*ompP5*, VAA clearly recognised terminal Gal*a*, in line with previous microcalorimetric studies revealing 20-fold increased affinity values for Gal*α*(1,4)Gal compared to lactose (Bharadwaj *et al.*, 1999). Intriguingly, the same lectin predominantly bound Gal*β* on the Rd LOS, demonstrating a surprising deficiency towards Gal*a*-recognition. Noteworthy, the presence of Gal*a* isoform in the Rd OS preparation was confirmed by heteronuclear correlation NMR experiments, corresponding to approximately 20% as previously reported (Risberg *et al.*, 1999). Consistent with our observation on VAA incapacity of Gal*a*-binding on the Rd LOS, this and other NTHi strains were negative for immunoblotting with the anti-globotriose monoclonal antibody 4C4 (High *et al.*, 1993) evidently pointing to unfavourable ligand

presentation. Among the strains tested for 4C4-reactivity also stand NTHi358 and 1566 (Martí-Lliteras *et al.*, 2011), yielding weaker fluorescence intensities for VAA-binding than NTHi375 in the microarray set-up, accordingly (section 4.1.8).

As galabiose recognition by VAA was feasible on the NTHi375-derived OS as shown by STD-NMR, a model structure was created using PyMOL (v. 1.6) and UCSF Chimera (v. 1.11) software, through manual superimposition. The *a*Gal(1,4)*β*Gal- disaccharide or the globotriose unit was superimposed onto the same atomic positions as for the Galp in the VAA crystal structure (Niwa *et al.*, 2003), Fig. 5.1, *a, b*.

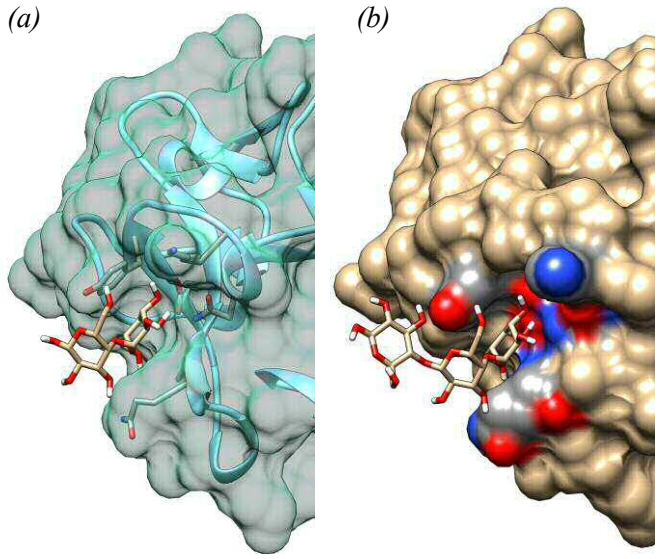


Fig. 5.1 Represantation of binding of Gal*α*(1,4)Gal and globotriose moieties to the 2γ carbohydrate-binding site of VAA. Representations of Gal*α*(1,4)Gal *(a)* and globotriose *(b)* in the same binding groove, 2γ, suggest a possible interaction mode with VAA (PDB: 1OQL) and illustrate that stronger contacts are projected only for the terminal *α*-Gal*p* moiety, furtherly supporting the acquired STD data for VAA-OS from NTHi375Δ*ompP5*. Aminoacid residues in *(b)* are coloured by element (red for oxygen and blue for nitrogen).

At protein concentration that low as 8 μM, VAA is still expected to form a dimer, and consequently spatial restrictions would apply to the accessibility of its Trp-binding site (1α, around Trp38) as previously demonstrated (Jiménez *et al.*, 2005; Jiménez *et al.*, 2006). Thus, the 2γ domain engaged around Tyr249 is the central glycan-binding site under our working conditions. Qualitatively, the suggested binding mode is compatible with the STD-deduced epitope mapping (Fig. 4.17 and Table 4.6). Both galabiose and globotriose seem to perfectly fit in the shallow groove-shaped binding site of VAA, but contacts established solely seem to include the CH-π interaction of the terminal Gal*a* H1/H3/H5 with the Tyr aromatic ring and bidentate H-bonding (Fig. 4.18), as

previously discussed. No direct contacts with Gal I seem to be possible, with the reducing end (Glc II) of the globotriose moiety, clearly staying outside the binding groove and heading towards the solvent, unable to establish contacts with the receptor (Fig. 5.1, *b*).

5.2 Conserved glycosignature for THP-1-derived exosomes

It is currently technically challenging to obtain pure EV fractions, free from non-vesicular components for functional studies, what makes substantial the common demand for establishment of a minimal set of biochemical, biophysical and functional parameters of their characterization (Lötvall *et al.*, 2014). In order to classify EVs according to particular glycosylation patterns they may carry, the prerequisite of previous physical characterisation was evident (Rupert *et al.*, 2017). EV preparations derived from THP-1 cell culture were submitted to ultracentrifugation and filtration fractionation, resulting in discrete EV populations of standard protein-to-lipid ratio, as suggested (Osteikoetxea *et al.*, 2015). Moreover, higher lipid order was evident for exosome preparations when compared to bigger-size elements, MVs, as function of the enrichment of exosomes in saturated phospholipids (Subra *et al.*, 2007) and their relevant curvature (de Jesus *et al.*, 2013; Kastelowitz and Yin, 2014).

Requirement for validation of the arrays with anti-CD63 antibody or Annexin V in parallel to the lectin binding assays, not simultaneously using different laser channels as done for bacterial microarrays, brought out the need for discovery of probes and markers beyond the phosphatidylserine and caspase molecules. Noteworthy, concerns about universality of these markers have risen, given their detection in preparations of different origin in an inconsistent manner (Jimenez *et al.*, 2003; Connor *et al.*, 2010; Yoshioka *et al.*, 2013).

Upon set-up of the EV-based microarrays, the study of glycosylation patterns of THP-1-derived extracellular vesicles by means of microarrays and lectin screening revealed conserved profiles, rich in Man, GlcNAc and Gal for exosomes. Regarding relevant abundance of these residues', our results for the THP-1-derived EV subpopulations were in line with previous observations for cancer cell-lines, including ovary (Escrevente *et al.*, 2011) and another six solid and liquid tumours (Batista *et al.*, 2011). Despite their diverse origin, exosomes isolated from healthy donors' blood were also

recognised from different lectins, the snowdrop lectin (GNA) attracting again the attention. All evidence comprised, it is primarily suggested that exosomes are enriched for oligomannose structures on their surface glycoproteins and, GNA specifically binds them. Besides exosomes, virions of hepatitis C (Tullis *et al.*, 2009) and HIV are also known to expose glycoproteins bearing high-mannose structures (Botos and Wlodawer, 2005; Hu *et al.*, 2007). In addition, previous work has provided strong evidence that HIV-1 and T-cell-derived EVs traffic out of host T-cells in a similar way, their glycosignatures being cell-line dependent yet sharing the same oligomannose epitope (Krishnamoorthy *et al.*, 2009). These results argue for a common origin and support the so-called Trojan exosome hypothesis (Gould *et al.*, 2003; Hirabayashi, 2009). In this context, antiviral lectins *i.e.* cyanovirin (Botos *et al.*, 2002), binding both virus and EV particles raise concern about targeting these glycans for therapeutics as suggested (Hu *et al.*, 2007), yet potentially explaining why eliciting a protective antiviral immune response is difficult.

In translational medicine a therapeutic hemofiltration approach emerged from the same observation, and was previously applied as adjuvant therapy permitting hepatitis C viral cargo reduction in dialysis patients, also having a remarkable impact in improving patient responses to the existent pharmaceutical remedies (Tullis *et al.*, 2009). A similar approach was later suggested under the brandname Haemopurifier®, based on a GNA-affinity-based plasmapheresis cartridge for exosome depletion during hemodialysis (Marleau *et al.*, 2012). Since the conserved oligomannose pattern is specifically recognised from GNA, its use as capturer guided the QCM sensors development for the study of exosomes. Indeed, GNA-derivatised chips were able to selectively capture exosomes contrary to MVs in flow, thus constituting a novel approach for efficient vesicle detection. The possibility of *in-flow* capture of exosomes even in body fluids paves the way for the establishment of novel, non-invasive biomarker detectors for clinical application, and is well inside the capacity of QCM instruments (Tombelli, 2012). This trend could also be exploited for designing new isolation protocols based on lectin affinity chromatography, distinguishing among EV subpopulations.

As for the kinetic analysis achieved directly on the surface of exosomes by QCM, further aspects of lectin domain organisation due to be considered. Even though data fit was proficient on the basis of a one-set-of-sites model for ConA and PNA, this was not the case for GNA. Usually, if a 1:1 model does not yield a proper fit, the existence of

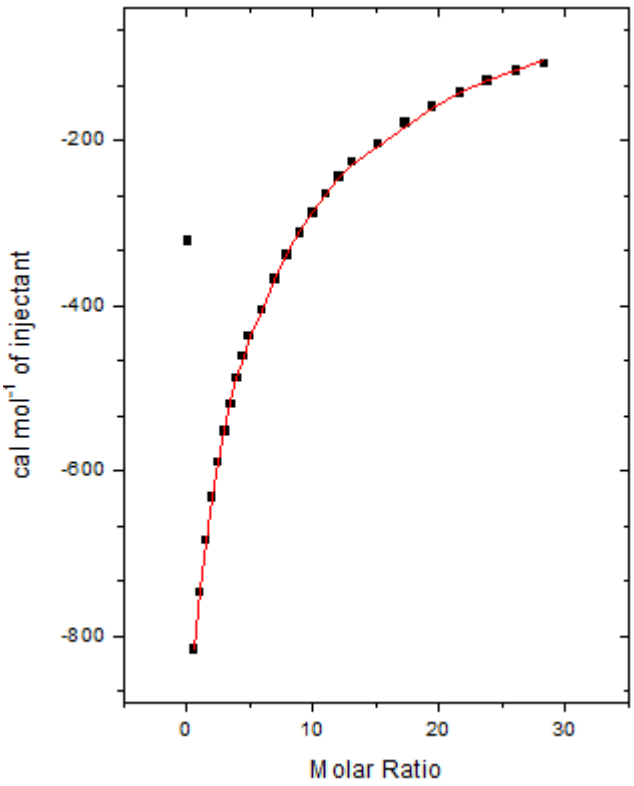


Fig. 5.2 Calorimetric titration of GNA with α-Me-diMan (■) at 22 °C. GNA was loaded in the calorimeter cell at 300 μM in HBS, pH 7.4, and the ligand was injected in a stepwise fashion. Continuous lines correspond to data fit to a two-set-of-sites model.

different docking sites for the lectin on the bacterial surface, or even ligand clustering, can be inferred. In this case however, the application of a 2:1 model was furtherly requested by the proper binding site arrangement in the lectin quaternary structure. Indeed, for this lectin not all sites are spaced at maximum distance, as occurs in the case of ConA. Instead, each monomer contains three subdomains, in close distance of the dimer interface as suggested by its crystallographic structure (Hester *et al.*, 1995; Hester and Wright, 1996). We have furtherly confirmed these observations in solution by isothermal titration calorimetry of GNA with *α*-Me-diMan. The equilibrium association constant for ligand binding at site 1 is nearly an order of magnitude higher than at site 2 (K_{a1} =8 × 10^{-4} M^{-1}; K_{a2} = 6.9 × 10^{-3} M^{-1}). Binding to site 2 had a 7-fold higher enthalpy variation. Noteworthy, GNA was reported to bind mono- and dimannosides at distinct sites (Hester and Wright, 1996). Interestingly, cyanovirin-N also contains two sets of sites in solution, shown to be responsible for multivalent binding between the lectin and the high-mannose chains, each site targeting a different part of the epitope (Shenoy *et al.*, 2002). Altogether, the observations point to specific binding of GNA on the THP-1 exosomes' surface, resulting in tight association and affinities in the submicromolar range (section 4.2.5).

5.3 Evaluation of glyconanoscaffolds of potential anti-adhesive application

Evidently, the microarray set-up permitted the evaluation of potent multivalent inhibitors and even enabled the comparison of their capacity, in the submicromolar range. Sensible fluorescent detection added an important advantage and along with miniaturisation and low quantity requirement, the approach proved out to be a valuable alternative to more tedious, time- and sample-consuming protocols previously established (Köhn *et al.*, 2004; Hartmann *et al.*, 2012; Barras *et al.*, 2013). Importantly, the approach also exhibited great reproducibility: control of bacterial viability and FimH stability is essential and clearly contributing to this direction. Previously reported repeatability deficits i.e. in the yeast aggregation assay (Firon *et al.*, 1983) may be possibly attributed to uncontrolled bacterial growth conditions, hereby proved to detrimentally affect the outcome (section 4.3.1).

Indeed, during the last couple of decades it has been considered that biofilm production, synthesis of capsular material and flagella motility is influenced by population density in several Gram-positive but also -negative bacteria, including *E. coli* (Miller and Bassler, 2001; Sperandio *et al.*, 2002). These "quorum sensing" bacteria produce and release chemical signal molecules (auto-inducers) that increase in concentration as a function of cell density. The detection of a minimal threshold stimulatory concentration of an auto-inducer leads to an alteration in gene expression. Interestingly, an auto-inducer called stationary phase-specific sigma factor (RpoS) was shown to have a negative regulatory role (probably indirect), at both the *fimA* and the *fimB* promoters, thus making a key contribution to the switching of cells between fimbriate and afimbriate states (Dove *et al.*, 1997). Further complex regulatory mechanisms were established regarding bacterial fimbriation through evolution. As relevant example, a drop in the intracellular concentration of cAMP (3′,5′-cyclic adenosine monophosphate) due to altered physiological conditions (e.g. growth in presence of glucose) increases the percentage of fimbriated cells in the bacterial population (Müller *et al.*, 2009). Accordingly, our rich, LB2YT medium used during *E. coli* UTI89 cultivation (section 3.1), could be furtherly promoting fimbriation.

As for the glycoside cluster effect as observed by ITC in the case of ConA and mannofullerenes (1) and (2) several observations may be done. Of note, the latter were found to be superior ConA inhibitors over silica nanoparticles, with almost two orders

of magnitude enhanced relative inhibitory potential per mannose residue. ITC results suggested binding of approximately 7 ConA monomer units per mannose-derivatised dodecafullerene (section 4.3.2.2.3). Calculation of the relative area corresponding to each mannose residue yielded *ca.* 38 nm^2 – considering the glycofullerene as a sphere of radio r = 6 nm (mean value for fullerenes bearing 2-4 PEG repeats in linker) and total area $A = 4 \times \pi \times r^2 = 452\ nm^2$. This relative area per mannose corresponds to an imaginable sphere of *ca.* 2 nm radio surrounding each residue. At this point, if one considers the distance between two adjacent ConA sites in the tetramer to be of 6.5 nm (Derewenda *et al.*, 1989), this would suggest a radio of 4.6 nm per tetramer and 1.15 nm ConA monomer. If all moieties are spherical, the accommodation of 12 monomers per fullerene seems rather impossible due to steric hindrance. What is more, the distance between vicinal ConA sites in the tetramer (6.5 nm) would not permit the establishment of contacts with adjacent mannose residues.

Thus, the importance of the linker length and flexibility in multivalent systems becomes evident. Most likely, the capacity of pegylated linkers to form a strong core resembling the SAMs (Harder *et al.*, 1998) – when it is distance-permitted – conditions the final availability of grafted sugars and their correct presentation, along with the hydrodynamic properties of the fullerenes. In this context, computational studies on the valency and behaviour of different classes of nanoscaffolds in the solvated state will probably be enlightening.

The high density of immobilised ConA on the chip could be responsible for a "bind and roll" effect (Munoz *et al.*, 2009), of particular relevance in the reliable evaluation of binding efficiencies in surface-based multivalent carbohydrate recognition. Evidently, this effect should apply for both fullerenes since the same chip was used for assessment. However, the linker length and core packing still conditions the final availability of mannose-residues for recognition. Further SPR experiments using chips of ConA immobilisation minimum 10-fold lower could contribute in rejecting bind and roll effects.

5.4 Perspectives

Despite the common trends in carbohydrate composition and display on the surface of diverse bacterial strains (Herget *et al.*, 2008; Sahly *et al.*, 2008), precise availability of glycosylation patterns for recognition by a given lectin may substantially differ.

Undoubtedly, studies on isolated macromolecules alone lead to the acquisition of solid information describing carbohydrate-lectin interactions. For example, our STD-NMR results illustrate how this ligand-based approach may confidently reveal the epitope for lectins at atomic level. Nonetheless, in these approaches the natural presentation and accessibility of glycans along with possible cooperativity of neighbour molecules are not considered. This aspect is of crucial relevance in the case of intrinsically weak binders as the lectins, whose affinity is controlled in multiple levels and are prone to synergistic interactions with third-party elements.

Thus, holistic methods that monitor the events directly at the surface may contribute to greater mechanistic understanding. In this direction, the developed bacteria-based microarrays served to gain information on the availability of certain motifs for endogenous lectin recognition and provided semi-quantitative data on binding avidity. These clues were validated and complemented by QCM kinetic analyses, providing the apparent affinity values and rate constants. Importantly, this combined approach operates directly at the surface thus closely mimicking the real bacteria–receptor interplay (Kalograiaki *et al.*, 2016).

Unravelling the binding patterns of lectins of the innate immune system by means of the newly developed platforms will serve as a starting point for further functional analyses regarding bacterial resistance and subversion of host lectins. Many bacteria bear extended LPS structures that are essential in shielding more vulnerable sites in the LPS inner core as suggested in this Dissertation and elsewhere (Matsuura, 2013; Clark *et al.*, 2016). Evidence points to mechanisms of these immunodominant molecules variation among others, in order to efficiently evade a first-line mucosal innate immune defense (Simon *et al.*, 1980; Finlay and McFadden, 2006). Of note, besides nontypeable *Haemophilus influenzae* (Weiser *et al.*, 1989) and *Klebsiella pneumoniae* here studied, several opportunistic respiratory pathogens such as *Neisseria* and *Pseudomonas* species, *Bordetella pertussis* and *Moraxella catarrhalis* are long used as models of antigenic variation (Henderson *et al.*, 1999), using a multitude of controlling element concepts and emphasizing why a vaccine to these organisms has not been successful yet. It is anticipated that high-throughput typification of Gram-negative clinical isolates and comparison of their virulence phenotypes and inter-strain glycosignatures with their relevant recognition by endogenous lectins will aid in the establishment of clinical correlations. Typification of bacterial strains by lectin fingerprinting based on both

binding pattern and binding affinity, could thereby add a further level of discrimination between strains. Importantly, the method is not limited to Gram-negative bacteria but also to Gram-positive or other pathogens as yeast, since the identification of high-affinity strain-selective binders may facilitate the development and optimization of lectin-based biosensors for their detection in clinical, environmental, and food samples.

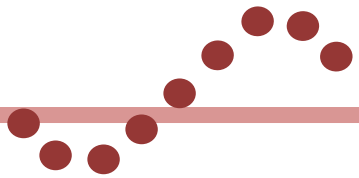
In sections 4.2.3-4, the availability of docking points for lectins on the EV surface was furtherly confirmed by the QCM vesicle-based chips, again allowing for the calculation of kinetic parameters in the glycans´ native milieu. Of note, GNA worked as a first-rate, selective capturer for exosomes also in the straight-forward in-flow capture approach, and was well recommended by the microarray-based lectin screening. Of general importance, the biosensor approach presented in this Dissertation could be tailored in order to address a variety of EV types and soluble molecules, e.g. aberrant glycoproteins, as potential biomarkers even directly in body fluids, diagnosis fully supported by the current state-of-the-art. In each case, previous glycophenotyping using the high-throughput microarray platforms may reveal other affinity agents of choice to better exploit the approach potential in diagnostics and clinical practice.

Finally, with respect to the microarray-assisted screening of anti-adhesive compounds, the approach introduced herein represents a powerful alternative for compound evaluation, also addressing the multiple needs arising from the use of live bacteria. The potential of the selected glycan-presenting nanoscaffolds as adhesion inhibitors was analysed by non-linear regression of several valid data points, and noticeably exposed even slight differences in the inhibitory potential between glycofullerenes of similar features, well within the acceptable error limit of 10%. Essential discrepancies between the theoretical expectations and the actual “functional affinity” values obtained for the nanoscaffolds can be attributed to the multivalent effect, as previously discussed.

Importantly, the original set-up here presented also proved out to permit differentiation between the fimbriate and afimbriate states of pathogenic bacteria, suggesting a full range of promising applications in microbiology. An additional approach of added value in which we are currently working, involves capture of tripsinised bladder epithelial cells in microarrays and further evaluation of anti-adhesive compounds as effective inhibitors of bacterial adhesion. In parallel, host and *E. coli* interplay could definitely be monitored in real-time by adapting the assay for QCM studies and directly growing the

host cells in COP-1 sensor chips, as previously reported (Pei *et al.*, 2012; Peiris *et al.*, 2012).

All things considered, the results presented in this Dissertation intended to respond to the constant need of novel high-throughput tools that permit quantitative glycan profiling and interaction monitoring with their host counterparts. Undoubtedly, a lot remains to be done in terms of application of these technologies in biomedicine and their integration in clinical diagnosis and therapy.



CONCLUSIONS

1. Bacteria-based microarrays have proved to be efficient for lectin typing of carbohydrate structures on the surface of bacteria lacking capsular polysaccharides and O-antigen-bearing lipopolysaccharides, thereby broadening the applicability of the initially established approach. Using non-typeable *Haemophilus influenzae* (NTHi) as model bacterium, strain- and lectin-specific binding patterns have been observed for a panel of NTHi375 mutants as well as for six different NTHi clinical isolates, also providing semi-quantitative data on binding avidity. The results have evidenced the availability of Man/Glc, sialic acid, and Gal residues on the bacterial surface for lectin recognition.

2. New Quartz Crystal Microbalance (QCM) NTHi sensor chips have been developed and used for label-free kinetic analysis of bacteria−lectin interactions directly on the bacterial surface, yielding quantitative data on binding affinity and kinetic parameters. Data obtained for the binding of the Gal-specific *Ricinus communis* agglutinin (RCA) revealed the formation of extremely stable complexes, while the binding profiles of wheat germ agglutinin and concanavalin A, which recognize sialic acid and Man/Glc residues, respectively, indicated the presence of more than one set of sites for these lectins, with noticeably faster dissociation rates compared to RCA.

3. Bacteria- and lipooligosaccharide (LOS)-based microarrays have revealed that RCA and the highly homologous Gal-specific agglutinin from *Viscum album* (VAA) recognize different ligands on the NTHi surface. In addition, STD-NMR experiments have served to define the recognized LOS epitopes, which are lectin- and oligosaccharide-specific. Only VAA binds strongly to the α-Gal-terminated LOS of NTHi375, whereas RCA shows a remarkable preference for the β-Gal-terminated LOS of NTHi strain Rd.

4. In line with the glycophenotyping results, binding to the NTHi arrays of three lectins of the innate immune system with Man/Glc-, sialic acid-, and Gal-binding specificity, *i.e.* SP-D, Siglec-14, and human galectin-8, has been observed, providing the first experimental evidence for direct binding of SP-D to NTHi. Moreover, comparison of the behaviour of selected galectins belonging to the three structural subgroups of this lectin family has unveiled striking differences in NTHi recognition, even when comparing the two carbohydrate-binding domains of tandem-repeat-type galectins. In particular, LOS truncation had detrimental effects on the binding to NTHi

of human galectins 4 and 8, pointing to LOS serving as ligand. The binding of galectin-8 was found to be carbohydrate-mediated, while for galectin-4 ionic interactions apparently drive recognition, red-ox conditions also playing a noticeable role.

5. Distinctive glycosylation profiles for EV subpopulations have been unveiled using novel EV microarrays and lectin screening. Pioneer vesicle-based QCM sensor chips have confirmed the binding trends observed for selected lectins in the microarray set-up and provided apparent affinity and kinetic parameters. In addition, a straight-forward approach for in-flow selective exosome detection has been developed, paving the way for non-invasive high-throughput screening of glycosignatures transported through biological fluids.

6. A microarray set-up for assessment of the functionality of *E. coli* FimH adhesin has been successfully established. The approach, involving analysis of the binding of fluorescently-labelled bacteria to mannan-printed arrays, has been found to be efficient for discriminating between *E. coli* fimbriate (FimH-expressing) and afimbriate states.

7. Using ConA as model lectin, the potential of designer microarrays for testing mannose-bearing silica nanoparticles and dodecafullerenes as inhibitors of lectin binding has been demonstrated, SPR and ITC analyses of ConA-inhibitor interactions serving to confirm the microarray-derived results. Mannofullerenes, found to be superior ConA inhibitors over silica nanoparticles, also proved to be efficient for preventing binding of FimH-expressing *E. coli* to mannan-printed arrays, with up to 18-fold enhancement of the relative inhibitory potential per mannose residue. The precise structure and length of the spacer used for linking mannose to the fullerene core has a noticeable impact on compound activity. To the best of our knowledge, this is the first report on a microarray approach enabling the screening of anti-adhesive compounds using live bacteria.

References

- Abe M., Suzuki O., Tasaki K., Tominaga K. and Wakasa H. (1996) Analysis of lectin binding properties on human Burkitt's lymphoma cell lines that show high spontaneous metastasis to distant organs in SCID mice: The binding sites for soybean agglutinin lectin masked by sialylation are closely associated with metastatic lymphoma cells. *Pathology International* **46**: 977-983.
- Abellán Flos M. (2015). Synthesis and functionalization of nanoplatforms targeting bacterial adhesion and multivalent enzyme inhibition. PhD Thesis, Université de Namur.
- Adak A. K., Boley J. W., Lyvers D. P., Chiu G. T., Low P. S., Reifenberger R. and Wei A. (2013) Label-free detection of *Staphylococcus aureus* captured on immutable ligand arrays. *ACS Applied Materials & Interfaces* **5**: 6404-6411.
- Adibekian A., Stallforth P., Hecht M.-L., Werz D. B., Gagneux P. and Seeberger P. H. (2011) Comparative bioinformatics analysis of the mammalian and bacterial glycomes. *Chemical Science* **2**: 337-344.
- Agrawal A. and Murphy T. F. (2011) *Haemophilus influenzae* infections in the *H. influenzae* type b conjugate vaccine era. *Journal of Clinical Microbiology* **49**: 3728-3732.
- Ahmad N., Gabius H.-J., André S., Kaltner H., Sabesan S., Roy R., Liu B., Macaluso F. and Brewer C. F. (2004) Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates and forms heterogeneous cross-linked complexes. *Journal of Biological Chemistry* **279**: 10841-10847.
- Akira S., Uematsu S. and Takeuchi O. (2006) Pathogen recognition and innate immunity. *Cell* **124**: 783-801.
- Alvarez R. A. and Blixt O. (2006) Identification of ligand specificities for glycan binding proteins using glycan arrays. In *Methods in Enzymology*. F. Minoru, Ed.: Academic Press. pp 292-310.
- Amann R. I., Binder B. J., Olson R. J., Chisholm S. W., Devereux R. and Stahl D. A. (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Applied and Environmental Microbiology* **56**: 1919-1925.
- Amon R., Reuven E. M., Leviatan Ben-Arye S. and Padler-Karavani V. (2014) Glycans in immune recognition and response. *Carbohydrate Research* **389**: 115-122.
- Anderson H., Jönsson M., Vestling L., Lindberg U. and Aastrup T. (2007) Quartz crystal microbalance sensor design: I. Experimental study of sensor response and performance. *Sensors and Actuators B: Chemical* **123**: 27-34.
- André S., Kojima S., Gundel G., Russwurm R., Schratt X., Unverzagt C. and Gabius H.-J. (2006) Branching mode in complex-type triantennary N-glycans as regulatory element of their ligand properties. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1760**: 768-782.
- Andreu Z. and Yáñez-Mó M. (2014) Tetraspanins in extracellular vesicle formation and function. *Frontiers in Immunology* **5**: 442.
- Angata T., Ishii T., Motegi T., Oka R., Taylor R., Soto P., Chang Y.-C., Secundino I., Gao C.-X., Ohtsubo K., Kitazume S., Nizet V., Varki A., Gemma A., Kida K. and Taniguchi N. (2013) Loss of Siglec-14 reduces the risk of chronic obstructive pulmonary disease exacerbation. *Cellular and Molecular Life Sciences* **70**: 3199-3210.

- Asensio J. L., Ardá A., Cañada F. J. and Jiménez-Barbero J. (2013) Carbohydrate–aromatic interactions. *Accounts of Chemical Research* **46**: 946-954.
- Asensio J. L., Cañada F. J., Siebert H.-C., Laynez J., Poveda A., Nieto P. M., Soedjanaamadja U. M., Gabius H.-J. and Jiménez-Barbero J. (2000) Structural basis for chitin recognition by defense proteins: GlcNAc residues are bound in a multivalent fashion by extended binding sites in hevein domains. *Chemistry & Biology* **7**: 529-543.
- Asensio J. L., Espinosa J. F., Dietrich H., Cañada F. J., Schmidt R. R., Martín-Lomas M. M., André S., Gabius H.-J. and Jiménez-Barbero J. (1999) Bovine heart Galectin-1 selects a unique (syn) conformation of C-Lactose, a flexible lactose analogue. *Journal of the American Chemical Society* **121**: 8995-9000.
- Avery O. T., MacLeod C. M. and McCarty M. (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *The Journal of Experimental Medicine* **79**: 137-158.
- Baenziger J. U. (1985) The role of glycosylation in protein recognition. Warner-Lambert Parke-Davis Award lecture. *The American Journal of Pathology* **121**: 382-391.
- Baenziger J. U. and Fiete D. (1979) Structural determinants of Ricinus communis agglutinin and toxin specificity for oligosaccharides. *Journal of Biological Chemistry* **254**: 9795-9799.
- Bahrani-Mougeot F. K., Buckles E. L., Lockatell C. V., Hebel J. R., Johnson D. E., Tang C. M. and Donnenberg M. S. (2002) Type 1 fimbriae and extracellular polysaccharides are preeminent uropathogenic *Escherichia coli* virulence determinants in the murine urinary tract. *Molecular Microbiology* **45**: 1079-1093.
- Bain C. D., Troughton E. B., Tao Y. T., Evall J., Whitesides G. M. and Nuzzo R. G. (1989) Formation of monolayer films by the spontaneous assembly of organic thiols from solution onto gold. *Journal of the American Chemical Society* **111**: 321-335.
- Barbieri M. (2008) The mechanisms of evolution: natural selection and natural conventions. In *The Codes of Life: The Rules of Macroevolution*. M. Barbieri and J. Hoffmeyer, Eds.: Springer Netherlands. pp 15-35.
- Barenkamp S. J. and Leininger E. (1992) Cloning, expression, and DNA sequence analysis of genes encoding nontypeable *Haemophilus influenzae* high-molecular-weight surface-exposed proteins related to filamentous hemagglutinin of *Bordetella pertussis*. *Infection and Immunity* **60**: 1302-1313.
- Barras A., Martin F. A., Bande O., Baumann J.-S., Ghigo J.-M., Boukherroub R., Beloin C., Siriwardena A. and Szunerits S. (2013) Glycan-functionalized diamond nanoparticles as potent *E. coli* anti-adhesives. *Nanoscale* **5**: 2307-2316.
- Barton G. M. and Medzhitov R. (2003) Toll-Like Receptor signaling pathways. *Science* **300**: 1524-1525.
- Batista B. S., Eng W. S., Pilobello K. T., Hendricks-Muñoz K. D. and Mahal L. K. (2011) Identification of a conserved glycan signature for microvesicles. *Journal of proteome research* **10**: 4624-4633.
- Baum L. G., Garner O. B., Schaefer K. and Lee B. (2014) Microbe–host interactions are positively and negatively regulated by galectin–glycan interactions. *Frontiers in Immunology* **5**.

- Bax A. and Davis D. G. (1985) MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *Journal of Magnetic Resonance (1969)* **65**: 355-360.
- Bayón C., He N., Deir-Kaspar M., Blasco P., André S., Gabius H.-J., Rumbero Á., Jiménez-Barbero J., Fessner W.-D. and Hernáiz M. J. (2016) Direct enzymatic branch-end extension of glycocluster-presented glycans: An effective strategy for programming glycan bioactivity. *Chemistry – A European Journal* **23**: 1623-1633.
- Bennett H. S. (1963) Morphological aspects of extracellular polysaccharides. *Journal of Histochemistry & Cytochemistry* **11**: 14-23.
- Beuvery E. C., van Rossum F. and Nagel J. (1982) Comparison of the induction of immunoglobulin M and G antibodies in mice with purified pneumococcal type 3 and meningococcal group C polysaccharides and their protein conjugates. *Infection and Immunity* **37**: 15-22.
- Bharadwaj S., Kaltner H., Korchagina E. Y., Bovin N. V., Gabius H.-J. and Surolia A. (1999) Microcalorimetric indications for ligand binding as a function of the protein for galactoside-specific plant and avian lectins. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1472**: 191-196.
- Bhattacharyya L. and Brewer C. F. (1988) Lectin-carbohydrate interactions. Studies of the nature of hydrogen bonding between D-galactose and certain D-galactose-specific lectins, and between D-mannose and concanavalin A. *European Journal of Biochemistry* **176**: 207-212.
- Blixt O., Head S., Mondala T., Scanlan C., Huflejt M. E., Alvarez R., Bryan M. C., Fazio F., Calarese D., Stevens J., Razi N., Stevens D. J., Skehel J. J., van Die I., Burton D. R., Wilson I. A., Cummings R., Bovin N., Wong C.-H. and Paulson J. C. (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 17033-17038.
- Borrelli S., Hegedus O., Shaw D. H., Jansson P. E. and Lindberg A. A. (1995) The tetrasaccharide L-alpha-D-heptose1-->2-L-alpha-D-heptose1--> 3-L-alpha-D-heptose1-->(3-deoxy-D-manno-octulosonic acid) and phosphate in lipid A define the conserved epitope in *Haemophilus lipopolysaccharides* recognized by a monoclonal antibody. *Infection and Immunity* **63**: 3683-3692.
- Botos I., O'Keefe B. R., Shenoy S. R., Cartner L. K., Ratner D. M., Seeberger P. H., Boyd M. R. and Wlodawer A. (2002) Structures of the complexes of a potent anti-HIV protein cyanovirin-N and high mannose oligosaccharides. *Journal of Biological Chemistry* **277**: 34336-34342.
- Botos I. and Wlodawer A. (2005) Proteins that bind high-mannose sugars of the HIV envelope. *Progress in Biophysics and Molecular Biology* **88**: 233-282.
- Bottazzi B., Inforzato A., Messa M., Barbagallo M., Magrini E., Garlanda C. and Mantovani A. (2016) The pentraxins PTX3 and SAP in innate immunity, regulation of inflammation and tissue remodelling. *Journal of Hepatology* **64**: 1416-1427.
- Bottom V. E. (1982) *Introduction to quartz crystal unit design*, Van Nostrand Reinhold.
- Bouckaert J., Berglund J., Schembri M., De Genst E., Cools L., Wuhrer M., Hung C.-S., Pinkner J., Slättegård R., Zavialov A., Choudhury D., Langermann S., Hultgren S. J., Wyns L., Klemm P., Oscarson S., Knight S. D. and De Greve H. (2005) Receptor binding studies disclose a novel class of high-affinity

inhibitors of the Escherichia coli FimH adhesin. *Molecular Microbiology* **55**: 441-455.
- Bouckaert J., Dewallef Y., Poortmans F., Wyns L. and Loris R. (2000) The structural features of concanavalin A governing non-proline peptide isomerization. *Journal of Biological Chemistry* **275**: 19778-19787.
- Bouchet V., Hood D. W., Li J., Brisson J.-R., Randle G. A., Martin A., Li Z., Goldstein R., Schweda E. K. H., Pelton S. I., Richards J. C. and Moxon E. R. (2003) Host-derived sialic acid is incorporated into *Haemophilus influenzae* lipopolysaccharide and is a major virulence factor in experimental otitis media. *Proceedings of the National Academy of Sciences* **100**: 8898-8903.
- Boukouris S. and Mathivanan S. (2015) Exosomes in bodily fluids are a highly stable resource of disease biomarkers. *PROTEOMICS – Clinical Applications* **9**: 358-367.
- Boyd W. C. (1954) The proteins of immune reactions. In *The Proteins*. H. Neurath and K. Bailey, Eds.: Academic Press, New York. pp 756–844.
- Boyd W. C. and Shapleigh E. (1954) Antigenic relations of blood group antigens as suggested by tests with lectins. *The Journal of Immunology* **73**: 226-231.
- Bradford M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* **72**: 248-254.
- Brewer C. F., Miceli M. C. and Baum L. G. (2002) Clusters, bundles, arrays and lattices: novel mechanisms for lectin–saccharide-mediated cellular interactions. *Current Opinion in Structural Biology* **12**: 616-623.
- Buddecke E. (2009) Proteoglycans. In *The Sugar Code. Fundamentals of Glycosciences*. H. J. Gabius, Ed. Weinheim, Germany: Wiley-VCH. pp 199-216.
- Burroughs M. H., Chang Y. S., Gage D. A. and Tuomanen E. I. (1993) Composition of the peptidoglycan of *Haemophilus influenzae*. *Journal of Biological Chemistry* **268**: 11594-11598.
- Buzamet E. (2012). Divergencias en las propiedades moleculares y modo de unión a ligandos de diferentes galectinas humanas. PhD, CSIC - Instituto de Química Física Rocasolano (IQFR) and Universidad Autónoma de Madrid.
- Calarese D. A., Lee H.-K., Huang C.-Y., Best M. D., Astronomo R. D., Stanfield R. L., Katinger H., Burton D. R., Wong C.-H. and Wilson I. A. (2005) Dissection of the carbohydrate specificity of the broadly neutralizing anti-HIV-1 antibody 2G12. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 13372-13377.
- Campanero-Rhodes M. A., Llobet E., Bengoechea J. A. and Solis D. (2015) Bacteria microarrays as sensitive tools for exploring pathogen surface epitopes and recognition by host receptors. *RSC Advances* **5**: 7173-7181.
- Carlsson S., Öberg C. T., Carlsson M. C., Sundin A., Nilsson U. J., Smith D., Cummings R. D., Almkvist J., Karlsson A. and Leffler H. (2007) Affinity of galectin-8 and its carbohydrate recognition domains for ligands in solution and at the cell surface. *Glycobiology* **17**: 663-676.
- Carof M., Lebbar S. and Szabó L. (1987) Detection of 3-deoxy-2-octulosonic acid in thiobarbiturate-negative endotoxins. *Carbohydrate Research* **161**: C4-C7.
- Carver J. P. (1991) Experimental structure determination of oligosaccharides. *Current Opinion in Structural Biology* **1**: 716-720.

- Clark H. W., Mackay R.-M., Deadman M. E., Hood D. W., Madsen J., Moxon E. R., Townsend J. P., Reid K. B. M., Ahmed A., Shaw A. J., Greenhough T. J. and Shrive A. K. (2016) Crystal structure of a complex of surfactant protein D (SP-D) and *Haemophilus influenzae* lipopolysaccharide reveals shielding of core structures in SP-D-resistant strains. *Infection and Immunity* **84**: 1585-1592.
- Cobb B. A. and Kasper D. L. (2005) Coming of age: carbohydrates and immunity. *European Journal of Immunology* **35**: 352-356.
- Cody A. J., Field D., Feil E. J., Stringer S., Deadman M. E., Tsolaki A. G., Gratz B., Bouchet V., Goldstein R., Hood D. W. and Moxon E. R. (2003) High rates of recombination in otitis media isolates of non-typeable Haemophilus influenzae. *Infection, Genetics and Evolution* **3**: 57-66.
- Collinge D. B., Kragh K. M., Mikkelsen J. D., Nielsen K. K., Rasmussen U. and Vad K. (1993) Plant chitinases. *The Plant Journal* **3**: 31-40.
- Compain P., Decroocq C., Iehl J., Holler M., Hazelard D., Mena Barragán T., Ortiz Mellet C. and Nierengarten J.-F. (2010) Glycosidase inhibition with fullerene iminosugar balls: A dramatic multivalent effect. *Angewandte Chemie* **122**: 5889-5892.
- Connor D. E., Exner T., Ma D. D. F. and Joseph J. E. (2010) The majority of circulating platelet-derived microparticles fail to bind annexin V, lack phospholipid-dependent procoagulant activity and demonstrate greater expression of glycoprotein Ib. *Thrombosis and Haemostasis* **103**: 1044-1052.
- Cooper D. N. W. (2002) Galectinomics: finding themes in complexity. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1572**: 209-231.
- Crocker P. R., Paulson J. C. and Varki A. (2007) Siglecs and their roles in the immune system. *Nature Reviews Immunology* **7**: 255-266.
- Crocker P. R. and Varki A. (2001) Siglecs in the immune system. *Immunology* **103**: 137-145.
- Crouch E. (1998) Structure, biologic properties, and expression of surfactant protein D (SP-D). *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1408**: 278-289.
- Crouch E. (2000) Surfactant protein-D and pulmonary host defense. *Respiratory Research* **1**: 93 - 108.
- Crouch E., Hartshorn K. and Ofek I. (2000) Collectins and pulmonary innate immunity. *Immunological Reviews* **173**: 52-65.
- Crouch E., Persson A., Chang D. and Heuser J. (1994) Molecular structure of pulmonary surfactant protein D (SP-D). *Journal of Biological Chemistry* **269**: 17311-17319.
- Cuatrecasas P. (1970) Protein purification by affinity chromatography: derivatizations of agarose and polyacrylamide beads. *Journal of Biological Chemistry* **245**: 3059-3065.
- Chabre Y. M. and Roy R. (2013) Multivalent glycoconjugate syntheses and applications using aromatic scaffolds. *Chemical Society Reviews* **42**: 4657-4708.
- Chang Y.-C. and Nizet V. (2014) The interplay between Siglecs and sialylated pathogens. *Glycobiology* **24**: 818-825.
- Chen H.-Y., Weng I. C., Hong M.-H. and Liu F.-T. (2014) Galectins as bacterial sensors in the host innate response. *Current Opinion in Microbiology* **17**: 75-81.
- Chien Y.-Y., Jan M.-D., Adak A. K., Tzeng H.-C., Lin Y.-P., Chen Y.-J., Wang K.-T., Chen C.-T., Chen C.-C. and Lin C.-C. (2008) Globotriose-functionalized

gold nanoparticles as multivalent probes for shiga-like toxin. *ChemBioChem* **9**: 1100-1109.

- Childs R. A., Palma A. S., Wharton S., Matrosovich T., Liu Y., Chai W., Campanero-Rhodes M. A., Zhang Y., Eickmann M., Kiso M., Hay A., Matrosovich M. and Feizi T. (2009) Receptor-binding specificity of pandemic influenza A (H1N1) 2009 virus determined by carbohydrate microarray. *Nature Biotechnology* **27**: 797-799.
- Choi D.-S., Kim D.-K., Kim Y.-K. and Gho Y. S. (2013) Proteomics, transcriptomics and lipidomics of exosomes and ectosomes. *Proteomics* **13**: 1554-1571.
- Choudhury D., Thompson A., Stojanoff V., Langermann S., Pinkner J., Hultgren S. J. and Knight S. D. (1999) Structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli*. *Science* **285**: 1061-1066.
- Chui D., Sellakumar G., Green R. S., Sutton-Smith M., McQuistan T., Marek K. W., Morris H. R., Dell A. and Marth J. D. (2001) Genetic remodeling of protein glycosylation in vivo induces autoimmune disease. *Proceedings of the National Academy of Sciences* **98**: 1142-1147.
- Dachary-Prigent J., Freyssinet J., Pasquet J., Carron J. and Nurden A. (1993) Annexin V as a probe of aminophospholipid exposure and platelet membrane vesiculation: a flow cytometry study showing a role for free sulfhydryl groups. *Blood* **81**: 2554-2565.
- Dalziel M., Crispin M., Scanlan C. N., Zitzmann N. and Dwek R. A. (2014) Emerging principles for the therapeutic exploitation of glycosylation. *Science* **343**.
- Dam T. K. and Brewer C. F. (2008) Effects of clustered epitopes in multivalent ligand−receptor Interactions. *Biochemistry* **47**: 8470-8476.
- Dam T. K. and Brewer C. F. (2010a) Lectins as pattern recognition molecules: The effects of epitope density in innate immunity. *Glycobiology* **20**: 270-279.
- Dam T. K. and Brewer C. F. (2010b) Multivalent lectin-carbohydrate interactions: energetics and mechanisms of binding. In *Advances in Carbohydrate Chemistry and Biochemistry*. H. Derek, Ed.: Academic Press. pp 139-164.
- Darkes M. M. and Plosker G. (2002) Pneumococcal conjugate vaccine (Prevnar™1; PNCRM7): A review of its use in the prevention of *Streptococcus pneumoniae* infection. *Pediatric Drugs* **4**: 609-630.
- Davicino R. C., Eliçabe R. J., Di Genaro M. S. and Rabinovich G. A. (2011) Coupling pathogen recognition to innate immunity through glycan-dependent mechanisms. *International Immunopharmacology* **11**: 1457-1463.
- Dawid S., Barenkamp S. J. and St. Geme J. W. (1999) Variation in expression of the *Haemophilus influenzae* HMW adhesins: A prokaryotic system reminiscent of eukaryotes. *Proceedings of the National Academy of Sciences* **96**: 1077-1082.
- De Crescenzo G., Boucher C., Durocher Y. and Jolicoeur M. (2008) Kinetic characterization by surface plasmon resonance-based biosensors: principle and emerging trends. *Cellular and Molecular Bioengineering* **1**: 204-215.
- De Chiara M., Hood D., Muzzi A., Pickard D. J., Perkins T., Pizza M., Dougan G., Rappuoli R., Moxon E. R., Soriani M. and Donati C. (2014) Genome sequencing of disease and carriage isolates of nontypeable *Haemophilus influenzae* identifies discrete population structure. *Proceedings of the National Academy of Sciences* **111**: 5439-5444.

- de Jesus A. J., Kastelowitz N. and Yin H. (2013) Changes in lipid density induce membrane curvature. *RSC Advances* **3**: 13622-13625.
- Derewenda Z., Yariv J., Helliwell J. R., Kalb A. J., Dodson E. J., Papiz M. Z., Wan T. and Campbell J. (1989) The structure of the saccharide-binding site of concanavalin A. *The EMBO Journal* **8**: 2189-2193.
- Domenech M., Pedrero-Vega E., Prieto A. and García E. (2016) Evidence of the presence of nucleic acids and β-glucan in the matrix of non-typeable Haemophilus influenzae in vitro biofilms. *Scientific Reports* **6**: 36424.
- Dove S. L., Smith S. G. J. and Dorman C. J. (1997) Control of *Escherichia coli* type 1 fimbrial gene expression in stationary phase: a negative role for RpoS. *Molecular and General Genetics MGG* **254**: 13-20.
- Dube D. H. and Bertozzi C. R. (2005) Glycans in cancer and inflammation - potential for therapeutics and diagnostics. *Nature Reviews Drug Discovery* **4**: 477-488.
- Dubois M., Gilles K., Hamilton J. K., Rebers P. A. and Smith F. (1951) A colorimetric method for the determination of sugars. *Nature* **168**: 167-167.
- Durka M., Buffet K., Iehl J., Holler M., Nierengarten J.-F., Taganna J., Bouckaert J. and Vincent S. P. (2011) The functional valency of dodecamannosylated fullerenes with *Escherichia coli* FimH-towards novel bacterial antiadhesives. *Chemical Communications* **47**: 1321-1323.
- Earl L. A., Bi S. and Baum L. G. (2011) Galectin multimerization and lattice formation are regulated by linker region structure. *Glycobiology* **21**: 6-12.
- Eischeid A. C. (2011) SYTO dyes and EvaGreen outperform SYBR Green in real-time PCR. *BMC Research Notes* **4**: 263-263.
- El Biari K. (2014). Estudio por RMN del reconocimiento molecular carbohidrato/proteína con discriminación diferencial de anómeros en equilibrio. PhD, CSIC - Centro de Investigaciones Biológicas (CIB) and Universidad Complutense de Madrid.
- Escaich S. (2008) Antivirulence as a new antibacterial approach for chemotherapy. *Current Opinion in Chemical Biology* **12**: 400-408.
- Escrevente C., Keller S., Altevogt P. and Costa J. (2011) Interaction and uptake of exosomes by ovarian cancer cells. *BMC Cancer* **11**: 108.
- Esko J. D., Kimata K. and Lindahl U. (2009) Proteoglycans and sulfated glycosaminoglycans. In *Essentials of Glycobiology*. A. Varki, R. D. Cummings, J. D. Esko et al., Eds. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press. pp 229-248.
- Euba B., Moleres J., Viadas C., Ruiz de los Mozos I., Valle J., Bengoechea J. A. and Garmendia J. (2015) Relative contribution of P5 and Hap surface proteins to nontypable *Haemophilus influenzae* interplay with the host upper and lower airways. *PLoS ONE* **10**: e0123154.
- Feizi T. (1985) Carbohydrate antigens in human cancer. *Cancer Surveys* **4**: 245–269.
- Feizi T., Fazio F., Chai W. and Wong C.-H. (2003) Carbohydrate microarrays — a new set of technologies at the frontiers of glycomics. *Current Opinion in Structural Biology* **13**: 637-645.
- Fernández-Alonso M. d. C., Cañada F. J., Jiménez-Barbero J. and Cuevas G. (2005) Molecular recognition of saccharides by proteins. Insights on the origin of the carbohydrate−aromatic interactions. *Journal of the American Chemical Society* **127**: 7379-7386.

- Finlay B. B. and McFadden G. (2006) Anti-Immunology: Evasion of the host immune system by bacterial and viral pathogens. *Cell* **124**: 767-782.
- Firon N., Ofek I. and Sharon N. (1983) Carbohydrate specificity of the surface lectins of *Escherichia coli*, *Klebsiella pneumoniae*, and *Salmonella typhimurium*. *Carbohydrate research* **120**: 235-249.
- Fleischmann R., Adams M., White O., Clayton R., Kirkness E., Kerlavage A., Bult C., Tomb J., Dougherty B., Merrick J. and al. e. (1995) Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science* **269**: 496-512.
- Fortini M. E. (2001) Sugar-coated pathways for developmental patterning. *Nature Cell Biology* **3**: E229-E231.
- Fox K. L., Cox A. D., Gilbert M., Wakarchuk W. W., Li J., Makepeace K., Richards J. C., Moxon E. R. and Hood D. W. (2006) Identification of a bifunctional lipopolysaccharide sialyltransferase in *Haemophilus influenzae*: incorporation of disialic acid. *Journal of Biological Chemistry* **281**: 40024-40032.
- Freire E., Mayorga O. L. and Straume M. (1990) Isothermal titration calorimetry. *Analytical Chemistry* **62**: 950A-959A.
- Frescatada-Rosa M., Stanislas T., Backues S. K., Reichardt I., Men S., Boutté Y., Jürgens G., Moritz T., Bednarek S. Y. and Grebe M. (2014) High lipid order of Arabidopsis cell-plate membranes mediated by sterol and dynamin-related protein 1A function. *The Plant Journal* **80**: 745-757.
- Freudenberg M. A. and Galanos C. (1990) Bacterial lipopolysaccharides: structure, metabolism and mechanisms of action. *International Reviews of Immunology* **6**: 207-221.
- Freyer M. W. and Lewis E. A. (2008) Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. *Methods Cell Biol* **84**: 79-113.
- Fujita T. (2002) Evolution of the lectin-complement pathway and its role in innate immunity. *Nature Reviews in Immunology* **2**: 346-353.
- Fujitani N., Furukawa J.-I., Araki K., Fujioka T., Takegawa Y., Piao J., Nishioka T., Tamura T., Nikaido T., Ito M., Nakamura Y. and Shinohara Y. (2013) Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers. *Proceedings of the National Academy of Sciences* **110**: 2105-2110.
- Fukiya S., Mizoguchi H., Tobe T. and Mori H. (2004) Extensive genomic diversity in pathogenic *Escherichia coli* and *Shigella* strains revealed by comparative genomic hybridization microarray. *Journal of Bacteriology* **186**: 3911-3921.
- Fusetti F., Pijning T., Kalk K. H., Bos E. and Dijkstra B. W. (2003) Crystal structure and carbohydrate-binding properties of the human cartilage glycoprotein-39. *Journal of Biological Chemistry* **278**: 37753-37760.
- Gabius H.-J. (2000) Biological information transfer beyond the genetic code: the Sugar Code. *Naturwissenschaften* **87**: 108-121.
- Gabius H.-J. (2009) Ca2+ : mastermind and active player for lectin activity. In *The Sugar Code. Fundamentals of Glycosciences*. H. J. Gabius, Ed. Weinheim, Germany: Wiley-VCH. pp 269-278.

- Gabius H.-J., André S., Jiménez-Barbero J., Romero A. and Solís D. (2011) From lectin structure to functional glycomics: principles of the Sugar Code. *Trends in Biochemical Sciences* **36**: 298-313.
- Gabius H.-J., Siebert H.-C., André S., Jiménez-Barbero J. and Rüdiger H. (2004) Chemical biology of the Sugar Code. *ChemBioChem* **5**: 740-764.
- Gagneux P. and Varki A. (1999) Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* **9**: 747-755.
- Garmendia J., Viadas C., Calatayud L., Mell J. C., Martí-Lliteras P., Euba B., Llobet E., Gil C., Bengoechea J. A., Redfield R. J. and Liñares J. (2014) Characterization of nontypable *Haemophilus influenzae* isolates recovered from adult patients with underlying chronic lung disease reveals genotypic and phenotypic traits associated with persistent infection. *PLoS ONE* **9**: e97020.
- Gawthorne J. A., Tan N. Y., Bailey U.-M., Davis M. R., Wong L. W., Naidu R., Fox K. L., Jennings M. P. and Schulz B. L. (2014) Selection against glycosylation sites in potential target proteins of the general HMWC N-glycosyltransferase in *Haemophilus influenzae*. *Biochemical and Biophysical Research Communications* **445**: 633-638.
- Geisler C. and Jarvis D. L. (2011) Letter to the Glyco-Forum: Effective glycoanalysis with Maackia amurensis lectins requires a clear understanding of their binding specificities. *Glycobiology* **21**: 988-993.
- Gerlach J. Q., Krüger A., Gallogly S., Hanley S. A., Hogan M. C., Ward C. J., Joshi L. and Griffin M. D. (2013) Surface Glycosylation Profiles of Urine Extracellular Vesicles. *PLoS ONE* **8**: e74801.
- Gibson B. W., Melaugh W., Phillips N. J., Apicella M. A., Campagnari A. A. and Griffiss J. M. (1993) Investigation of the structural heterogeneity of lipooligosaccharides from pathogenic *Haemophilus* and *Neisseria* species and of R-type lipopolysaccharides from *Salmonella typhimurium* by electrospray mass spectrometry. *Journal of Bacteriology* **175**: 2702-2712.
- Godula K. and Bertozzi C. R. (2012) Density variant glycan microarray for evaluating cross-linking of mucin-like glycoconjugates by lectins. *Journal of the American Chemical Society* **134**: 15732-15742.
- Goldstein I. J., Hollerman C. E. and Smith E. E. (1965) Protein-Carbohydrate Interaction. II. Inhibition studies on the interaction of Concanavalin A with polysaccharides. *Biochemistry* **4**: 876-883.
- Gomes J., Gomes-Alves P., Carvalho S., Peixoto C., Alves P., Altevogt P. and Costa J. (2015) Extracellular vesicles from ovarian carcinoma cells display specific glycosignatures. *Biomolecules* **5**: 1741.
- Gottlieb H. E., Kotlyar V. and Nudelman A. (1997) NMR chemical shifts of common laboratory solvents as trace impurities. *The Journal of Organic Chemistry* **62**: 7512-7515.
- Gouin S. G., Wellens A., Bouckaert J. and Kovensky J. (2009) Synthetic multimeric heptyl mannosides as potent antiadhesives of uropathogenic *Escherichia coli*. *ChemMedChem* **4**: 749-755.
- Gould S. J., Booth A. M. and Hildreth J. E. K. (2003) The Trojan exosome hypothesis. *Proceedings of the National Academy of Sciences* **100**: 10592-10597.
- Grass S., Buscher A. Z., Swords W. E., Apicella M. A., Barenkamp S. J., Ozchlewski N. and St Geme J. W. (2003) The *Haemophilus influenzae* HMW1 adhesin is glycosylated in a process that requires HMW1C and

phosphoglucomutase, an enzyme involved in lipooligosaccharide biosynthesis. *Molecular Microbiology* **48**: 737-751.

- Grass S., Lichti C. F., Townsend R. R., Gross J. and St. Geme J. W., III (2010a) The *Haemophilus influenzae* HMW1C protein is a glycosyltransferase that transfers hexose residues to asparagine sites in the HMW1 adhesin. *PLoS Pathog* **6**: e1000919.
- Grass S., Lichti C. F., Townsend R. R., Gross J. and St. Geme J. W., III (2010b) The *Haemophilus influenzae* HMW1C protein is a glycosyltransferase that transfers hexose residues to asparagine sites in the HMW1 adhesin. *PLoS Pathogens* **6**: e1000919.
- Green R. S., Stone E. L., Tenno M., Lehtonen E., Farquhar M. G. and Marth Jamey D. (2007) Mammalian N-glycan branching protects against innate immune self-recognition and inflammation in autoimmune disease pathogenesis. *Immunity* **27**: 308-320.
- Gronow S., Brabetz W., Lindner B. and Brade H. (2005) OpsX from *Haemophilus influenzae* represents a novel type of heptosyltransferase i in lipopolysaccharide biosynthesis. *Journal of Bacteriology* **187**: 6242-6247.
- Gross J., Grass S., Davis A. E., Gilmore-Erdmann P., Townsend R. R. and St. Geme J. W. (2008) The *Haemophilus influenzae* HMW1 adhesin is a glycoprotein with an unusual N-linked carbohydrate modification. *Journal of Biological Chemistry* **283**: 26010-26015.
- Gupta G. and Surolia A. (2007) Collectins: sentinels of innate immunity. *BioEssays* **29**: 452-464.
- Gupta G. S. (2012) Lectins: an overview. The animal lectin families. In *Animal Lectins: Form, Function and Clinical Applications*. G. S. Gupta, Ed.: Springer-Verlag Wien. pp 5-8.
- György B., Módos K., Pállinger É., Pálóczi K., Pásztói M., Misják P., Deli M. A., Sipos Á., Szalai A., Voszka I., Polgár A., Tóth K., Csete M., Nagy G., Gay S., Falus A., Kittel Á. and Buzás E. I. (2011a) Detection and isolation of cell-derived microparticles are compromised by protein complexes resulting from shared biophysical parameters. *Blood* **117**: e39-e48.
- György B., Szabó T. G., Pásztói M., Pál Z., Misják P., Aradi B., László V., Pállinger É., Pap E., Kittel Á., Nagy G., Falus A. and Buzás E. I. (2011b) Membrane vesicles, current state-of-the-art: emerging role of extracellular vesicles. *Cellular and Molecular Life Sciences* **68**: 2667-2688.
- Haan L. d. and Hirst T. R. (2004) Cholera toxin: A paradigm for multi-functional engagement of cellular mechanisms. *Molecular Membrane Biology* **21**: 77-92.
- Hakomori S.-I. (1985) Aberrant glycosylation in cancer cell membranes as focused on glycolipids: overview and perspectives. *Cancer Research* **45**: 2405-2414.
- Hakomori S.-I. (2002) Glycosylation defining cancer malignancy: New wine in an old bottle. *Proceedings of the National Academy of Sciences* **99**: 10231-10233.
- Hall D. R., Cann J. R. and Winzor D. J. (1996) Demonstration of an upper limit to the range of association rate constants amenable to study by biosensor technology based on surface plasmon resonance. *Analytical Biochemistry* **235**: 175-184.

- Hamouda H., Ullah M., Berger M., Sittinger M., Tauber R., Ringe J. and Blanchard V. (2013) N-Glycosylation Profile of undifferentiated and adipogenically differentiated human bone marrow mesenchymal stem cells: Towards a next generation of stem cell markers. *Stem Cells and Development* **22**: 3100-3113.
- Hao N., Neranon K., Ramström O. and Yan M. (2016) Glyconanomaterials for biosensing applications. *Biosensors and Bioelectronics* **76**: 113-130.
- Harder P., Grunze M., Dahint R., Whitesides G. M. and Laibinis P. E. (1998) Molecular conformation in oligo(ethylene glycol)-terminated self-assembled monolayers on gold and silver surfaces determines their ability to resist protein adsorption. *The Journal of Physical Chemistry B* **102**: 426-436.
- Hardman K. D. and Ainsworth C. F. (1972) Structure of concanavalin A at 2.4-Ang resolution. *Biochemistry* **11**: 4910-4919.
- Harrison A., Dyer D. W., Gillaspy A., Ray W. C., Mungur R., Carson M. B., Zhong H., Gipson J., Gipson M., Johnson L. S., Lewis L., Bakaletz L. O. and Munson R. S. (2005) Genomic sequence of an otitis media isolate of nontypeable *Haemophilus influenzae*: Comparative study with *H. influenzae* serotype d, strain KW20. *Journal of Bacteriology* **187**: 4627-4636.
- Hartmann M., Betz P., Sun Y., Gorb S. N., Lindhorst T. K. and Krueger A. (2012) Saccharide-modified nanodiamond conjugates for the efficient detection and removal of pathogenic bacteria. *Chemistry – A European Journal* **18**: 6485-6492.
- Hay E. D. (1991) Introductory remarks. In *Cell biology of the extracellular matrix*. New York: Springer Science+Business Media. pp 1-4.
- Heidecke C. D. and Lindhorst T. K. (2007) Iterative synthesis of spacered glycodendrons as oligomannoside mimetics and evaluation of their antiadhesive properties. *Chemistry – A European Journal* **13**: 9056-9067.
- Helander I. M., Lindner B., Brade H., Altmann K., Lindberg A. A., Rietschel E. T. and ZÄHringer U. (1988) Chemical structure of the lipopolysaccharide of *Haemophilus influenzae* strain I-69 Rd−/b+. *European Journal of Biochemistry* **177**: 483-492.
- Henderson I. R., Owen P. and Nataro J. P. (1999) Molecular switches — the ON and OFF of bacterial phase variation. *Molecular Microbiology* **33**: 919-932.
- Hennet T., Chui D., Paulson J. C. and Marth J. D. (1998) Immune regulation by the ST6Gal sialyltransferase. *Proceedings of the National Academy of Sciences* **95**: 4504-4509.
- Herget S., Toukach P., Ranzinger R., Hull W., Knirel Y. and von der Lieth C.-W. (2008) Statistical analysis of the Bacterial Carbohydrate Structure Data Base (BCSDB): characteristics and diversity of bacterial carbohydrates in comparison with mammalian glycans. *BMC Structural Biology* **8**: 35.
- Hernáiz M. J., De La Fuente J. M., Barrientos Á. G. and Penadés S. (2002) A model system mimicking glycosphingolipid clusters to quantify carbohydrate self interactions by surface plasmon resonance. *Angewandte Chemie International Edition* **41**: 1554-1557.
- Hester G., Kaku H., Goldstein I. J. and Wright C. S. (1995) Structure of mannose-specific snowdrop (Galanthus nivalis) lectin is representative of a new plant lectin family. *Nat Struct Mol Biol* **2**: 472-479.
- Hester G. and Wright C. S. (1996) The mannose-specific bulb lectin from *Galanthus nivalis* (snowdrop) binds mono- and dimannosides at distinct sites.

Structure analysis of refined complexes at 2.3 Å and 3.0 Å resolution. *Journal of Molecular Biology* **262**: 516-531.

- High N. J., Deadman M. E. and Moxon E. R. (1993) The role of a repetitive DNA motif (5′-CAAT-3′) in the variable expression of the *Haemophilus influenzae* lipopolysaccharide epitope *α*Gal(1–4)*β*Gal. *Molecular Microbiology* **9**: 1275-1282.
- Hirabayashi J. (2009) Glycome 'fingerprints' provide definitive clues to HIV origins. *Nat Chem Biol* **5**: 198-199.
- Hirabayashi J. and Kasai K.-i. (1993) The family of metazoan metal-independent β-galactoside-binding lectins: structure, function and molecular evolution. *Glycobiology* **3**: 297-304.
- Holgersson J., Gustafsson A. and Gaunitz S. (2009) Bacterial and viral lectins. In *The Sugar Code. Fundamentals of Glycosciences*. H.-J. Gabius, Ed. Weinheim, Germany: Wiley-VCH. pp 279-299.
- Homola J. (2003) Present and future of surface plasmon resonance biosensors. *Analytical and Bioanalytical Chemistry* **377**: 528-539.
- Homola J. (2008) Surface plasmon resonance sensors for detection of chemical and biological species. *Chemical Reviews* **108**: 462-493.
- Hood D. W., Cox A. D., Wakarchuk W. W., Schur M., Schweda E. K. H., Walsh S. L., Deadman M. E., Martin A., Moxon E. R. and Richards J. C. (2001) Genetic basis for expression of the major globotetraose-containing lipopolysaccharide from H. influenzae strain Rd (RM118). *Glycobiology* **11**: 957-967.
- Hood D. W., Deadman M. E., Cox A. D., Makepeace K., Martin A., Richards J. C. and Moxon E. R. (2004) Three genes, *lgtF*, *lic2C* and *lpsA*, have a primary role in determining the pattern of oligosaccharide extension from the inner core of *Haemophilus influenzae* LPS. *Microbiology* **150**: 2089-2097.
- Hood D. W., Makepeace K., Deadman M. E., Rest R. F., Thibault P., Martin A., Richards J. C. and Moxon E. R. (1999) Sialic acid in the lipopolysaccharide of *Haemophilus influenzae*: strain distribution, influence on serum resistance and structural characterization. *Molecular Microbiology* **33**: 679-692.
- Hopwood D. (1969) A comparison of the crosslinking abilities of glutaraldehyde, formaldehyde and α-hydroxyadipaldehyde with bovine serum albumin and casein. *Histochemie* **17**: 151-161.
- Hsu K.-L., Pilobello K. T. and Mahal L. K. (2006) Analyzing the dynamic bacterial glycome with a lectin microarray approach. *Nature Chemical Biology* **2**: 153-157.
- Hu Q., Mahmood N. and Shattock R. J. (2007) High-mannose-specific deglycosylation of HIV-1 gp120 induced by resistance to cyanovirin-N and the impact on antibody neutralization. *Virology* **368**: 145-154.
- Hudak J. E., Canham S. M. and Bertozzi C. R. (2014) Glycocalyx engineering reveals a Siglec-based mechanism for NK cell immunoevasion. *Nature Chemical Biology* **10**: 69-75.
- Huflejt M. E. and Leffler H. (2003) Galectin-4 in normal tissues and cancer. *Glycoconjugate Journal* **20**: 247-255.
- Hunter S. A. and Cochran J. R. (2016) Chapter Two - Cell-binding assays for determining the affinity of protein–protein interactions: Technologies and considerations. In *Methods in Enzymology*. L. P. Vincent, Ed.: Academic Press. pp 21-44.

- Iehl J. and Nierengarten J.-F. (2009) A click–click approach for the preparation of functionalized [5:1]-hexaadducts of C60. *Chemistry – A European Journal* **15**: 7306-7309.
- Iozzo R. V. (2005) Basement membrane proteoglycans: from cellar to ceiling. *Nature Reviews Molecular Cell Biology* **6**: 646-656.
- Ismail N., Wang Y., Dakhlallah D., Moldovan L., Agarwal K., Batte K., Shah P., Wisler J., Eubank T. D., Tridandapani S., Paulaitis M. E., Piper M. G. and Marsh C. B. (2013) Macrophage microvesicles induce macrophage differentiation and miR-223 transfer. *Blood* **121**: 984-995.
- Iwashkiw J. A., Vozza N. F., Kinsella R. L. and Feldman M. F. (2013) Pour some sugar on it: the expanding world of bacterial protein O-linked glycosylation. *Molecular Microbiology* **89**: 14-28.
- Jack D. L., Read R. C., Tenner A. J., Frosch M., Turner M. W. and Klein N. J. (2001) Mannose-binding lectin regulates the inflammatory response of human professional phagocytes to *Neisseria meningitidis* serogroup B. *Journal of Infectious Diseases* **184**: 1152-1162.
- Jaipuri F. A., Collet B. Y. M. and Pohl N. L. (2008) Synthesis and quantitative evaluation of glycero-D-manno-heptose binding to Concanavalin A by fluorous-tag assistance. *Angewandte Chemie International Edition* **47**: 1707-1710.
- Janshoff A., Galla H.-J. and Steinem C. (2000) Piezoelectric mass-sensing devices as biosensors—an alternative to optical biosensors? *Angewandte Chemie International Edition* **39**: 4004-4032.
- Jansson P.-E., Stenutz R. and Widmalm G. (2006) Sequence determination of oligosaccharides and regular polysaccharides using NMR spectroscopy and a novel Web-based version of the computer program casper. *Carbohydrate Research* **341**: 1003-1010.
- Jimenez J. J., Jy W., Mauro L. M., Soderland C., Horstman L. L. and Ahn Y. S. (2003) Endothelial cells release phenotypically and quantitatively distinct microparticles in activation and apoptosis. *Thrombosis Research* **109**: 175-180.
- Jiménez M., André S., Barillari C., Romero A., Rognan D., Gabius H.-J. and Solís D. (2008) Domain versatility in plant AB-toxins: Evidence for a local, pH-dependent rearrangement in the 2γ lectin site of the mistletoe lectin by applying ligand derivatives and modelling. *FEBS Letters* **582**: 2309-2312.
- Jiménez M., André S., Siebert H.-C., Gabius H.-J. and Solís D. (2006) AB-type lectin (toxin/agglutinin) from mistletoe: differences in affinity of the two galactoside-binding Trp/Tyr-sites and regulation of their functionality by monomer/dimer equilibrium. *Glycobiology* **16**: 926-937.
- Jiménez M., Sáiz J. L., André S., Gabius H.-J. and Solís D. (2005) Monomer/dimer equilibrium of the AB-type lectin from mistletoe enables combination of toxin/agglutinin activities in one protein: analysis of native and citraconylated proteins by ultracentrifugation/gel filtration and cell biological consequences of dimer destabilization. *Glycobiology* **15**: 1386-1395.
- Johannsmann D. (2008) Viscoelastic, mechanical, and dielectric measurements on complex samples with the quartz crystal microbalance. *Physical Chemistry Chemical Physics* **10**: 4516-4534.
- Johnson K. G. and Perry M. B. (1976) Improved techniques for the preparation of bacterial lipopolysaccharides. *Canadian Journal of Microbiology* **22**: 29-34.
- Jurcisek J., Greiner L., Watanabe H., Zaleski A., Apicella M. A. and Bakaletz L. O. (2005) Role of sialic acid and complex carbohydrate biosynthesis in biofilm

formation by nontypeable *Haemophilus influenzae* in the chinchilla middle ear. *Infection and Immunity* **73**: 3210-3218.
- Kabat E. A., Heidelberger M. and Bezer A. E. (1947) A study of the purification and properties of ricin. *Journal of Biological Chemistry* **168**: 629-639.
- Kalograiaki I., Euba B., Proverbio D., Campanero-Rhodes M. A., Aastrup T., Garmendia J. and Solís D. (2016) Combined Bacteria Microarray and Quartz Crystal Microbalance approach for exploring glycosignatures of nontypeable *Haemophilus influenzae* and recognition by host lectins. *Analytical Chemistry* **88**: 5950-5957.
- Kalra H., Adda C. G., Liem M., Ang C.-S., Mechler A., Simpson R. J., Hulett M. D. and Mathivanan S. (2013) Comparative proteomics evaluation of plasma exosome isolation techniques and assessment of the stability of exosomes in normal human blood plasma. *Proteomics* **13**: 3354-3364.
- Kalra H., Simpson R. J., Ji H., Aikawa E., Altevogt P., Askenase P., Bond V. C., Borràs F. E., Breakefield X., Budnik V., Buzas E., Camussi G., Clayton A., Cocucci E., Falcon-Perez J. M., Gabrielsson S., Gho Y. S., Gupta D., Harsha H. C., Hendrix A., Hill A. F., Inal J. M., Jenster G., Krämer-Albers E.-M., Lim S. K., Llorente A., Lötvall J., Marcilla A., Mincheva-Nilsson L., Nazarenko I., Nieuwland R., Nolte-'t Hoen E. N. M., Pandey A., Patel T., Piper M. G., Pluchino S., Prasad T. S. K., Rajendran L., Raposo G., Record M., Reid G. E., Sánchez-Madrid F., Schiffelers R. M., Siljander P., Stensballe A., Stoorvogel W., Taylor D., Thery C., Valadi H., van Balkom B. W. M., Vázquez J., Vidal M., Wauben M. H. M., Yáñez-Mó M., Zoeller M. and Mathivanan S. (2012) Vesiclepedia: a compendium for extracellular vesicles with continuous community annotation. *PLoS Biology* **10**: e1001450.
- Kaltner H., Solís D., Kopitz J., Lensch M., Lohr M., Manning J., Murnseer M., Schnolzer M., André S. and Saiz J. (2008) Prototype chicken galectins revisited: characterization of a third protein with distinctive hydrodynamic behaviour and expression pattern in organs of adult animals. *Biochemical Journal* **409**: 591-599.
- Karlsson K.-A. (1995) Microbial recognition of target-cell glycoconjugates. *Current Opinion in Structural Biology* **5**: 622-635.
- Kasper D. L., Paoletti L. C., Wessels M. R., Guttormsen H. K., Carey V. J., Jennings H. J. and Baker C. J. (1996) Immune response to type III group B streptococcal polysaccharide-tetanus toxoid conjugate vaccine. *The Journal of Clinical Investigation* **98**: 2308-2314.
- Kastelowitz N. and Yin H. (2014) Exosomes and microvesicles: Identification and targeting by particle size and lipid chemical probes. *ChemBioChem* **15**: 923-928.
- Keerthikumar S., Chisanga D., Ariyaratne D., Al Saffar H., Anand S., Zhao K., Samuel M., Pathan M., Jois M., Chilamkurti N., Gangoda L. and Mathivanan S. (2016) ExoCarta: A web-based compendium of exosomal cargo. *Journal of Molecular Biology* **428**: 688-692.
- Kelly D. F., Moxon E. R. and Pollard A. J. (2004) *Haemophilus influenzae* type b conjugate vaccines. *Immunology* **113**: 163-174.
- Kilcoyne M., Twomey M. E., Gerlach J. Q., Kane M., Moran A. P. and Joshi L. (2014) Campylobacter jejuni strain discrimination and temperature-dependent glycome expression profiling by lectin microarray. *Carbohydrate Research* **389**: 123-133.

- Kim D.-K., Kang B., Kim O. Y., Choi D.-s., Lee J., Kim S. R., Go G., Yoon Y. J., Kim J. H., Jang S. C., Park K.-S., Choi E.-J., Kim K. P., Desiderio D. M., Kim Y.-K., Lötvall J., Hwang D. and Gho Y. S. (2013) EVpedia: an integrated database of high-throughput data for systemic analyses of extracellular vesicles. *Journal of Extracellular Vesicles* **2**: 10.3402/jev.v3402i3400.20384.
- Kim S.-H., Turnbull J. and Guimond S. (2011) Extracellular matrix and cell signalling: the dynamic cooperation of integrin, proteoglycan and growth factor receptor. *Journal of Endocrinology* **209**: 139-151.
- Kishore U., Greenhough T. J., Waters P., Shrive A. K., Ghai R., Kamran M. F., Bernal A. L., Reid K. B. M., Madan T. and Chakraborty T. (2006) Surfactant proteins SP-A and SP-D: Structure, function and receptors. *Molecular Immunology* **43**: 1293-1315.
- Kline K. A., Fälker S., Dahlberg S., Normark S. and Henriques-Normark B. (2009) Bacterial Adhesins in host-microbe interactions. *Cell Host & Microbe* **5**: 580-592.
- Kohler S., Ullrich S., Richter U. and Schumacher U. (2009) E-/P-selectins and colon carcinoma metastasis: first in vivo evidence for their crucial role in a clinically relevant model of spontaneous metastasis formation in the lung. *British Journal of Cancer* **102**: 602-609.
- Köhn M., Benito J. M., Ortiz Mellet C., Lindhorst T. K. and García Fernández J. M. (2004) Functional evaluation of carbohydrate-centred glycoclusters by enzyme-linked lectin assay: Ligands for Concanavalin A. *ChemBioChem* **5**: 771-777.
- Kong N., Xie S., Zhou J., Menéndez M., Solís D., Park J., Proietti G., Ramström O. and Yan M. (2016) Catalyst-free cycloaddition reaction for the synthesis of glyconanoparticles. *ACS Applied Materials & Interfaces* **8**: 28136-28142.
- Kong N., Zhou J., Park J., Xie S., Ramström O. and Yan M. (2015) Quantitative fluorine NMR to determine carbohydrate density on glyconanomaterials synthesized from perfluorophenyl azide-functionalized silica nanoparticles by click reaction. *Analytical Chemistry* **87**: 9451-9458.
- Kopitz J. (2009) Glycolipids. In *The Sugar Code. Fundamentals of Glycosciences*. H. J. Gabius, Ed. Weinheim, Germany: Wiley-VCH. pp 177-197.
- Kožár T., André S., Uličný J. and Gabius H.-J. (2009) Three-dimensional aspects of the Sugar Code. In *The Sugar Code. Fundamentals of Glycosciences*. H. J. Gabius, Ed. Weinheim, Germany: Wiley-VCH. pp 15-29.
- Kretschmann E. and Raether H. (1968) Radiative decay of nonradiative surface plasmons excited by light. *Zeitschrift für Naturforschung A* **23**: 2135-2136.
- Krishnamoorthy L., Bess J. W., Preston A. B., Nagashima K. and Mahal L. K. (2009a) HIV-1 and microvesicles from T cells share a common glycome, arguing for a common origin. *Nat Chem Biol* **5**: 244-250.
- Krishnamoorthy L., Bess J. W., Preston A. B., Nagashima K. and Mahal L. K. (2009b) HIV-1 and microvesicles from T cells share a common glycome, arguing for a common origin. *Nature Chemical Biology* **5**: 244-250.
- Krogfelt K. A., Bergmans H. and Klemm P. (1990) Direct evidence that the FimH protein is the mannose-specific adhesin of *Escherichia coli* type 1 fimbriae. *Infection and Immunity* **58**: 1995-1998.
- Kuan S. F., Rust K. and Crouch E. (1992) Interactions of surfactant protein D with bacterial lipopolysaccharides. Surfactant protein D is an *Escherichia coli*-

binding protein in bronchoalveolar lavage. *The Journal of Clinical Investigation* **90**: 97-106.
- Kumar H., Kawai T. and Akira S. (2011a) Pathogen recognition by the innate immune system. *International Reviews in Immunology* **30**: 16-34.
- Kumar H., Kawai T. and Akira S. (2011b) Pathogen recognition by the innate immune system. *International Reviews of Immunology* **30**: 16-34.
- Kurokawa T., Tsuda M. and Sugino Y. (1976) Purification and characterization of a lectin from *Wistaria floribunda* seeds. *Journal of Biological Chemistry* **251**: 5686-5693.
- Laemmli U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685.
- Laine R. A. (2008) The information-storing potential of the Sugar Code. In *Glycosciences: Status and Perspectives*. H.-J. Gabius and S. Gabius, Eds. Weinheim, Germany: Wiley-VCH. pp 1-14.
- Lasky L. A. (1995) Selectin-carbohydrate interactions and the initiation of the inflammatory response. *Annual Review of Biochemistry* **64**: 113-140.
- Laurence D. J. R., Stevens U., Bettelheim R., Darcy D., Leese C., Turberville C., Alexander P., Johns E. W. and Neville A. M. (1972) Role of plasma carcinoembryonic antigen in diagnosis of gastrointestinal, mammary, and bronchial carcinoma. *The British Medical Journal* **3**: 605-609.
- Lee C.-H. and Tsai C.-M. (1999) Quantification of bacterial lipopolysaccharides by the Purpald assay: measuring formaldehyde generated from 2-keto-3-deoxyoctonate and heptose at the inner core by periodate oxidation. *Analytical Biochemistry* **267**: 161-168.
- Lee C.-S., Kim Y.-G., Joo H.-S. and Kim B.-G. (2004) Structural analysis of lipid A from *Escherichia coli* O157:H7:K− using thin-layer chromatography and ion-trap mass spectrometry. *Journal of Mass Spectrometry* **39**: 514-525.
- Lee R. T., Gabius H. J. and Lee Y. C. (1992) Ligand binding characteristics of the major mistletoe lectin. *Journal of Biological Chemistry* **267**: 23722-23727.
- Lee R. T. and Lee Y. C. (2000) Affinity enhancement by multivalent lectin–carbohydrate interaction. *Glycoconjugate Journal* **17**: 543-551.
- Lemieux R. U. (1989) Rhone-Poulenc Lecture. The origin of the specificity in the recognition of oligosaccharides by proteins. *Chemical Society Reviews* **18**: 347-374.
- Li X., Pei Y., Zhang R., Shuai Q., Wang F., Aastrup T. and Pei Z. (2013) A suspension-cell biosensor for real-time determination of binding kinetics of protein-carbohydrate interactions on cancer cell surfaces. *Chemical Communications* **49**: 9908-9910.
- Li X., Song S., Shuai Q., Pei Y., Aastrup T., Pei Y. and Pei Z. (2015) Real-time and label-free analysis of binding thermodynamics of carbohydrate-protein interactions on unfixed cancer cell surfaces using a QCM biosensor. *Scientific Reports* **5**: 14066.
- Liang Y., Eng W. S., Colquhoun D. R., Dinglasan R. R., Graham D. R. and Mahal L. K. (2014) Complex N-linked glycans serve as a determinant for exosome/microvesicle cargo recruitment. *Journal of Biological Chemistry* **289**: 32526-32537.
- Lin C.-C., Yeh Y.-C., Yang C.-Y., Chen G.-F., Chen Y.-C., Wu Y.-C. and Chen C.-C. (2003) Quantitative analysis of multivalent interactions of carbohydrate-

encapsulated gold nanoparticles with concanavalin A. *Chemical communications*: 2920-2921.

- Liu F.-T. and Rabinovich G. A. (2005) Galectins as modulators of tumour progression. *Nature Reviews in Cancer* **5**: 29-41.
- López-Lucendo M. F., Solís D., André S., Hirabayashi J., Kasai K.-i., Kaltner H., Gabius H.-J. and Romero A. (2004) Growth-regulatory human galectin-1: crystallographic characterisation of the structural changes induced by single-site mutations and their impact on the thermodynamics of ligand binding. *Journal of Molecular Biology* **343**: 957-970.
- Lötvall J., Hill A. F., Hochberg F., Buzás E. I., Di Vizio D., Gardiner C., Gho Y. S., Kurochkin I. V., Mathivanan S., Quesenberry P., Sahoo S., Tahara H., Wauben M. H., Witwer K. W. and Théry C. (2014a) Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the International Society for Extracellular Vesicles. *Journal of Extracellular Vesicles* **3**: 26913.
- Lötvall J., Hill A. F., Hochberg F., Buzás E. I., Di Vizio D., Gardiner C., Gho Y. S., Kurochkin I. V., Mathivanan S., Quesenberry P., Sahoo S., Tahara H., Wauben M. H., Witwer K. W. and Théry C. (2014b) Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the International Society for Extracellular Vesicles. *Journal of Extracellular Vesicles* **3**: 10.3402/jev.v3403.26913.
- Lowe C. R. (1985) An introduction to the concepts and technology of biosensors. *Biosensors* **1**: 3-16.
- Lowe J. B. and Marth J. D. (2003) A genetic approach to mammalian glycan function. *Annual Review of Biochemistry* **72**: 643-691.
- Lundquist J. J. and Toone E. J. (2002) The cluster glycoside effect. *Chemical Reviews* **102**: 555-578.
- Lütteke T. and Frank M. (2009) Synergy of computational and experimental methods in carbohydrate 3D structure determination and validation. In *Bioinformatics for Glycobiology and Glycomics*. C.-W. Von der Lieth, T. Lütteke and M. Frank, Eds. Chichester, UK: John Wiley & Sons, Ltd. pp 389-412.
- Llobet E., Campos M. A., Giménez P., Moranta D. and Bengoechea J. A. (2011) Analysis of the networks controlling the antimicrobial-peptide-dependent induction of *Klebsiella pneumoniae* virulence factors. *Infection and Immunity* **79**: 3718-3732.
- Macauley M. S., Crocker P. R. and Paulson J. C. (2014) Siglec-mediated regulation of immune cell function in disease. *Nature Reviews Immunology* **14**: 653-666.
- Macauley M. S. and Paulson J. C. (2014) Glyco-engineering 'super-self'. *Nature Chemical Biology* **10**: 7-8.
- Mann D. A., Kanai M., Maly D. J. and Kiessling L. L. (1998) Probing low affinity and multivalent interactions with surface plasmon resonance: Ligands for Concanavalin A. *Journal of the American Chemical Society* **120**: 10575-10582.
- Marleau A. M., Chen C.-S., Joyce J. A. and Tullis R. H. (2012) Exosome removal as a therapeutic adjuvant in cancer. *Journal of Translational Medicine* **10**: 134-134.

- Marth J. D. and Grewal P. K. (2008) Mammalian glycosylation in immunity. *Nature Reviews Immunology* **8**: 874-887.
- Martí-Lliteras P., López-Gómez A., Mauro S., Hood D. W., Viadas C., Calatayud L., Morey P., Servin A., Liñares J., Oliver A., Bengoechea J. A. and Garmendia J. (2011) Nontypable *Haemophilus influenzae* displays a prevalent surface structure molecular pattern in clinical isolates. *PLoS ONE* **6**: e21133.
- Marx K. A. (2003) Quartz crystal microbalance: a useful tool for studying thin polymer films and complex biomolecular systems at the solution−surface interface. *Biomacromolecules* **4**: 1099-1120.
- Mason J. T. and O'Leary T. J. (1991) Effects of formaldehyde fixation on protein secondary structure: a calorimetric and infrared spectroscopic investigation. *Journal of Histochemistry & Cytochemistry* **39**: 225-229.
- Masoud H., Moxon E. R., Martin A., Krajcarski D. and Richards J. C. (1997) Structure of the variable and conserved lipopolysaccharide oligosaccharide epitopes expressed by *Haemophilus influenzae* serotype b strain eagan. *Biochemistry* **36**: 2091-2103.
- Matsuura M. (2013) Structural modifications of bacterial lipopolysaccharide that facilitate gram-negative bacteria evasion of host innate immunity. *Frontiers in Immunology* **4**.
- Maughan H. and Redfield R. J. (2009) Tracing the evolution of competence in *Haemophilus influenzae*. *PLoS ONE* **4**: e5854.
- Mayer M. and Meyer B. (2001) Group epitope mapping by Saturation Transfer Difference NMR to identify segments of a ligand in direct contact with a protein receptor. *Journal of the American Chemical Society* **123**: 6108-6117.
- McGreal E. P., Martinez-Pomares L. and Gordon S. (2004) Divergent roles for C-type lectins expressed by cells of the innate immune system. *Molecular Immunology* **41**: 1109-1121.
- Medzhitov R. and Janeway C. A. (2002) Decoding the patterns of self and nonself by the innate immune system. *Science* **296**: 298-300.
- Melo S. A., Luecke L. B., Kahlert C., Fernandez A. F., Gammon S. T., Kaye J., LeBleu V. S., Mittendorf E. A., Weitz J., Rahbari N., Reissfelder C., Pilarsky C., Fraga M. F., Piwnica-Worms D. and Kalluri R. (2015) Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. *Nature* **523**: 177-182.
- Mell J. C., Sinha S., Balashov S., Viadas C., Grassa C. J., Ehrlich G. D., Nislow C., Redfield R. J. and Garmendia J. (2014) Complete genome sequence of *Haemophilus influenzae* strain 375 from the middle ear of a pediatric patient with otitis media. *Genome Announcements* **2**.
- Mell J. C., Viadas C., Moleres J., Sinha S., Fernández-Calvet A., Porsch E. A., St. Geme J. W., III, Nislow C., Redfield R. J. and Garmendia J. (2016a) Transformed recombinant enrichment profiling rapidly identifies HMW1 as an intracellular invasion locus in *Haemophilus influenzae*. *PLoS Pathogens* **12**: e1005576.
- Mell J. C., Viadas C., Moleres J., Sinha S., Fernández-Calvet A., Porsch E. A., St. Geme J. W., III, Nislow C., Redfield R. J. and Garmendia J. (2016b) Transformed recombinant enrichment profiling rapidly identifies HMW1 as an intracellular invasion locus in *Haemophilus influenzae*. *PLoS Pathog* **12**: e1005576.

- Mey A., Leffler H., Hmama Z., Normier G. and Revillard J. P. (1996) The animal lectin galectin-3 interacts with bacterial lipopolysaccharides via two independent sites. *The Journal of Immunology* **156**: 1572-1577.
- Miller M. B. and Bassler B. L. (2001) Quorum sensing in bacteria. *Annual Review of Microbiology* **55**: 165-199.
- Miller M. C., Ribeiro J. P., Roldós V., Martín-Santamaría S., Cañada F. J., Nesmelova I. A., André S., Pang M., Klyosov A. A., Baum L. G., Jiménez-Barbero J., Gabius H.-J. and Mayo K. H. (2011) Structural aspects of binding of α-linked digalactosides to human galectin-1. *Glycobiology* **21**: 1627-1641.
- Mirabella F. M. (1985) Internal reflection spectroscopy. *Applied Spectroscopy Reviews* **21**: 45-178.
- Mitchell S. W. (1861) Researches upon the venom of the rattlesnake: with an investigation of the anatomy and physiology of the organs concerned. In.: Smithsonian Institution. pp 89-90.
- Mittelbrunn M. and Sánchez-Madrid F. (2012) Intercellular communication: diverse structures for exchange of genetic information. *Nature Reviews Molecular Cell Biology* **13**: 328-335.
- Miyamoto N. and Bakaletz L. O. (1996) Selective adherence of non-typeable *Haemophilus influenzae* (NTHi) to mucus or epithelial cells in the chinchilla Eustachian tube and middle ear. *Microbial Pathogenesis* **21**: 343-356.
- Mo H., Winter H. C. and Goldstein I. J. (2000) Purification and characterization of a Neu5Acα2–6Galβ1–4Glc/GlcNAc-specific lectin from the fruiting body of the polypore mushroom *Polyporus squamosus*. *Journal of Biological Chemistry* **275**: 10623-10629.
- Monsigny M., Roche A.-C., Sene C., Maget-Dana R. and Delmotte F. (1980) Sugar-lectin interactions: how does wheat-germ agglutinin bind sialoglycoconjugates? *European Journal of Biochemistry* **104**: 147-153.
- Moran A. P., Prendergast M. M. and Appelmelk B. J. (1996) Molecular mimicry of host structures by bacterial lipopolysaccharides and its contribution to disease. *FEMS Immunology & Medical Microbiology* **16**: 105-115.
- Morey P., Cano V., Martí-Lliteras P., López-Gómez A., Regueiro V., Saus C., Bengoechea J. A. and Garmendia J. (2011) Evidence for a non-replicative intracellular stage of nontypable *Haemophilus influenzae* in epithelial cells. *Microbiology* **157**: 234-250.
- Morey P., Viadas C., Euba B., Hood D. W., Barberán M., Gil C., Grilló M. J., Bengoechea J. A. and Garmendia J. (2013) Relative contributions of lipooligosaccharide inner and outer core modifications to nontypeable *Haemophilus influenzae* pathogenesis. *Infection and Immunity* **81**: 4100-4111.
- Mulvey M. A. (2002) Adhesion and entry of uropathogenic *Escherichia coli*. *Cellular Microbiology* **4**: 257-271.
- Müller C. M., Åberg A., Straseviçiene J., Emődy L., Uhlin B. E. and Balsalobre C. (2009) Type 1 fimbriae, a colonization factor of uropathogenic *Escherichia coli,* are controlled by the metabolic sensor CRP-cAMP. *PLoS Pathogens* **5**: e1000303.
- Munoz E. M., Correa J., Fernandez-Megia E. and Riguera R. (2009) Probing the relevance of lectin clustering for the reliable evaluation of multivalent carbohydrate recognition. *Journal of the American Chemical Society* **131**: 17765-17767.

- Munoz E. M., Correa J., Riguera R. and Fernandez-Megia E. (2013) Real-Time evaluation of binding mechanisms in multivalent interactions: A surface plasmon resonance kinetic approach. *Journal of the American Chemical Society* **135**: 5966-5969.
- Musante L., Tataruch-Weinert D., Kerjaschki D., Henry M., Meleady P. and Holthofer H. (2017) Residual urinary extracellular vesicles in ultracentrifugation supernatants after hydrostatic filtration dialysis enrichment. *Journal of Extracellular Vesicles* **6**: 1267896.
- Naegeli A., Michaud G., Schubert M., Lin C.-W., Lizak C., Darbre T., Reymond J.-L. and Aebi M. (2014a) Substrate specificity of cytoplasmic *N*-glycosyltransferase. *Journal of Biological Chemistry* **289**: 24521-24532.
- Naegeli A., Neupert C., Fan Y.-Y., Lin C.-W., Poljak K., Papini A. M., Schwarz F. and Aebi M. (2014b) Molecular analysis of an alternative *N*-glycosylation machinery by functional transfer from *Actinobacillus pleuropneumoniae* to *Escherichia coli*. *The Journal of Biological Chemistry* **289**: 2170-2179.
- Nesmelova I. V., Ermakova E., Daragan V. A., Pang M., Menéndez M., Lagartera L., Solís D., Baum L. G. and Mayo K. H. (2010) Lactose binding to galectin-1 modulates structural dynamics, increases conformational entropy, and occurs with apparent negative cooperativity. *Journal of Molecular Biology* **397**: 1209-1230.
- Nierengarten I. and Nierengarten J.-F. (2014) Fullerene sugar balls: A new class of biologically active fullerene derivatives. *Chemistry – An Asian Journal* **9**: 1436-1444.
- Nierengarten J.-F., Iehl J., Oerthel V., Holler M., Illescas B. M., Munoz A., Martin N., Rojo J., Sanchez-Navarro M., Cecioni S., Vidal S., Buffet K., Durka M. and Vincent S. P. (2010) Fullerene sugar balls. *Chemical Communications* **46**: 3860-3862.
- Niwa H., Tonevitsky A. G., Agapov I. I., Saward S., Pfüller U. and Palmer R. A. (2003) Crystal structure at 3 Å of mistletoe lectin I, a dimeric type-II ribosome-inactivating protein, complexed with galactose. *European Journal of Biochemistry* **270**: 2739-2749.
- Nothaft H. and Szymanski C. M. (2010) Protein glycosylation in bacteria: sweeter than ever. *Nature Reviews Microbiology* **8**: 765-778.
- Nsimba-Lubaki M., Peumans W. J. and Allen A. K. (1986) Isolation and characterization of glycoprotein lectins from the bark of three species of elder, *Sambucus ebulus*, *S. nigra* and *S. racemosa*. *Planta* **168**: 113-118.
- Oda Y., Herrmann J., Gitt M. A., Turck C. W., Burlingame A. L., Barondes S. H. and Leffler H. (1993) Soluble lactose-binding lectin from rat intestine with two different carbohydrate-binding domains in the same peptide chain. *Journal of Biological Chemistry* **268**: 5929-5939.
- Ofek I. and Beachey E. H. (1978) Mannose binding and epithelial cell adherence of *Escherichia coli*. *Infection and Immunity* **22**: 247-254.
- Ofek I., Mesika A., Kalina M., Keisari Y., Podschun R., Sahly H., Chang D., McGregor D. and Crouch E. (2001) Surfactant Protein D enhances phagocytosis and killing of unencapsulated phase variants of *Klebsiella pneumoniae*. *Infection and Immunity* **69**: 24-33.
- Ofek I., Mirelman D. and Sharon N. (1977) Adherence of *Escherichia coli* to human mucosal cells mediated by mannose receptors. *Nature* **265**: 623-625.

- Ohtsubo K. and Marth J. D. (2006) Glycosylation in cellular mechanisms of health and disease. *Cell* **126**: 855-867.
- Olsnes S., Sandvig K., Eiklid K. and Pihl A. (1978) Properties and action mechanism of the toxic lectin modeccin: Interaction with cell lines resistant to modeccin, abrin, and ricin. *Journal of Supramolecular Structure* **9**: 15-25.
- Osborn M. J. (1963) Studies on the Gram-negative cell wall, I. Evidence for the role of 2-keto-3-deoxyoctonate in the lipopolysaccharide of *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences* **50**: 499-506.
- Osteikoetxea X., Balogh A., Szabó-Taylor K., Németh A., Szabó T. G., Pálóczi K., Sódar B., Kittel Á., György B., Pállinger É., Matkó J. and Buzás E. I. (2015) Improved characterization of EV preparations based on protein to lipid ratio and lipid properties. *PLoS ONE* **10**: e0121184.
- Otto A. (1968) Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. *Zeitschrift für Physik* **216**: 398-410.
- Owen D. M., Rentero C., Magenau A., Abu-Siniyeh A. and Gaus K. (2012) Quantitative imaging of membrane lipid order in cells and organisms. *Nat. Protocols* **7**: 24-35.
- Palma A. S., Feizi T., Zhang Y., Stoll M. S., Lawson A. M., Díaz-Rodríguez E., Campanero-Rhodes M. A., Costa J., Gordon S., Brown G. D. and Chai W. (2006) Ligands for the β-Glucan receptor, Dectin-1, assigned using "designer" microarrays of oligosaccharide probes (neoglycolipids) generated from glucan polysaccharides. *Journal of Biological Chemistry* **281**: 5771-5779.
- Parasassi T., Krasnowska E. K., Bagatolli L. and Gratton E. (1998) Laurdan and Prodan as polarity-sensitive fluorescent membrane probes. *Journal of Fluorescence* **8**: 365-373.
- Patsos G. and Corfield A. (2009) O-Glycosylation: structural diversity and functions. In *The Sugar Code. Fundamentals of Glycosciences.* H. J. Gabius, Ed. Weinheim, Germany: Wiley-VCH. pp 111-137.
- Pei Z., Saint-Guirons J., Käck C., Ingemarsson B. and Aastrup T. (2012) Real-time analysis of the carbohydrates on cell surfaces using a QCM biosensor: a lectin-based approach. *Biosensors and Bioelectronics* **35**: 200-205.
- Peiris D., Markiv A., Curley G. P. and Dwek M. V. (2012) A novel approach to determining the affinity of protein–carbohydrate interactions employing adherent cancer cells grown on a biosensor surface. *Biosensors and Bioelectronics* **35**: 160-166.
- Peracaula R., Barrabés S., Sarrats A., Rudd P. M. and de Llorens R. (2008) Altered glycosylation in tumours focused to cancer diagnosis. *Disease markers* **25**: 207-218.
- Persson A., Chang D. and Crouch E. (1990) Surfactant protein D is a divalent cation-dependent carbohydrate-binding protein. *Journal of Biological Chemistry* **265**: 5755-5760.
- Peterson P. K. and Quie P. G. (1981a) Bacterial surface components and the pathogenesis of infectious diseases. *Annual Reviews in Medicine* **32**: 29-43.
- Peterson P. K. and Quie P. G. (1981b) Bacterial surface components and the pathogenesis of infectious diseases. *Annual Review of Medicine* **32**: 29-43.
- Peumans W. J. and Van Damme E. J. M. (1995) Lectins as plant defense proteins. *Plant Physiology* **109**: 347-352.

- Phillips N. J., Apicella M. A., Griffiss J. M. and Gibson B. W. (1992) Structural characterization of the cell surface lipooligosaccharides from a nontypable strain of *Haemophilus influenzae*. *Biochemistry* **31**: 4515-4526.
- Pinho S. S. and Reis C. A. (2015) Glycosylation in cancer: mechanisms and clinical implications. *Nature Reviews Cancer* **15**: 540-555.
- Pizarro-Cerdá J. and Cossart P. (2006) Bacterial adhesion and entry into host cells. *Cell* **124**: 715-727.
- Podder S. K., Surolia A. and Bachhawat B. K. (1974) On the specificity of carbohydrate-lectin recognition. *European Journal of Biochemistry* **44**: 151-160.
- Ponnuraj K., Bowden M. G., Davis S., Gurusiddappa S., Moore D., Choe D., Xu Y., Hook M. and Narayana S. V. L. (2003) A "dock, lock, and latch" structural model for a staphylococcal adhesin binding to fibrinogen. *Cell* **115**: 217-228.
- Preston A., Mandrell R. E., Gibson B. W. and Apicella M. A. (1996) The lipooligosaccharides of pathogenic Gram-negative bacteria. *Critical Reviews in Microbiology* **22**: 139-180.
- Quattroni P., Li Y., Lucchesi D., Lucas S., Hood D. W., Herrmann M., Gabius H.-J., Tang C. M. and Exley R. M. (2012) Galectin-3 binds *Neisseria meningitidis* and increases interaction with phagocytic cells. *Cellular microbiology* **14**: 1657-1675.
- Qureshi N., Mascagni P., Ribi E. and Takayama K. (1985) Monophosphoryl lipid A obtained from lipopolysaccharides of *Salmonella minnesota* R595. Purification of the dimethyl derivative by high performance liquid chromatography and complete structural determination. *Journal of Biological Chemistry* **260**: 5271-5278.
- Rabe M., Verdes D. and Seeger S. (2011) Understanding protein adsorption phenomena at solid surfaces. *Advances in Colloid and Interface Science* **162**: 87-106.
- Rabinovich G. A. and Toscano M. A. (2009) Turning 'sweet' on immunity: galectin-glycan interactions in immune tolerance and inflammation. *Nature Reviews in Immunology* **9**: 338-352.
- Rabinovich G. A., van Kooyk Y. and Cobb B. A. (2012) Glycobiology of immune responses. *Annals of the New York Academy of Sciences* **1253**: 1-15.
- Rademacher T. W., Parekh R. B. and Dwek R. A. (1988) Glycobiology. *Annual Review of Biochemistry* **57**: 785-838.
- Raedler A., Schreiber S. and Batsakis J. G. (1988) Analysis of differentiation and transformation of cells by lectins. *CRC Critical Reviews in Clinical Laboratory Sciences* **26**: 153-193.
- Raether H. (1988) *Surface plasmons on smooth and rough surfaces and on gratings*. New York, Springer Berlin Heidelberg.
- Raetz C. R. H., Reynolds C. M., Trent M. S. and Bishop R. E. (2007) Lipid A modification systems in Gram-negative bacteria. *Annual Review of Biochemistry* **76**: 295-329.
- Rao V. K., Krasan G. P., Hendrixson D. R., Dawid S. and St. Geme J. W. (1999) Molecular determinants of the pathogenesis of disease due to non-typable *Haemophilus influenzae*. *FEMS Microbiology Reviews* **23**: 99-129.
- Raposo G. and Stoorvogel W. (2013) Extracellular vesicles: Exosomes, microvesicles, and friends. *The Journal of Cell Biology* **200**: 373-383.

- Reddy M. S., Bernstein J. M., Murphy T. F. and Faden H. S. (1996) Binding between outer membrane proteins of nontypeable *Haemophilus influenzae* and human nasopharyngeal mucin. *Infection and Immunity* **64**: 1477-1479.
- Rees D. A. (1972) Shapely polysaccharides. The eighth Colworth medal lecture. *The Biochemical Journal* **126**: 257-273.
- Reichardt N. C., Martin-Lomas M. and Penades S. (2013) Glyconanotechnology. *Chemical Society Reviews* **42**: 4358-4376.
- Reid C. W., Fulton K. M. and Twine S. M. (2010) Never take candy from a stranger: the role of the bacterial glycome in host–pathogen interactions. *Future Microbiology* **5**: 267-288.
- Reinhardt A., Wehle M., Geissner A., Crouch E. C., Kang Y., Yang Y., Anish C., Santer M. and Seeberger P. H. (2016) Structure binding relationship of human surfactant protein D and various lipopolysaccharide inner core structures. *Journal of Structural Biology* **195**: 387-395.
- Remaut H. and Waksman G. (2004) Structural biology of bacterial pathogenesis. *Current Opinion in Structural Biology* **14**: 161-170.
- Rini J. M. (1995) Lectin structure. *Annual Review of Biophysics and Biomolecular Structure* **24**: 551-577.
- Risberg A., Masoud H., Martin A., Richards J. C., Moxon E. R. and Schweda E. K. H. (1999) Structural analysis of the lipopolysaccharide oligosaccharide epitopes expressed by a capsule-deficient strain of *Haemophilus influenzae* Rd. *European Journal of Biochemistry* **261**: 171-180.
- Rivera-Sagredo A., Jiménez-Barbero J., Martín-Lomas M., Solís D. and Díaz-Mauriño T. (1992) Studies of the molecular recognition of synthetic methyl *β*-lactoside analogues by *Ricinus communis* agglutinin. *Carbohydrate Research* **232**: 207-226.
- Roth J., Zuber C., Park S., Jang I., Lee Y., Kysela K., Le Fourn V., Santimaria R., Guhl B. and Cho J. (2010) Protein N-glycosylation, protein folding, and protein quality control. *Molecules and Cells* **30**: 497-506.
- Roy R., Page D., Figueroa Perez S. and Verez Bencomo V. (1998) Effect of shape, size, and valency of multivalent mannosides on their binding properties to phytohemagglutinins. *Glycoconjugate Journal* **15**: 251-263.
- Ruiz F. M., Scholz B. A., Buzamet E., Kopitz J., André S., Menéndez M., Romero A., Solís D. and Gabius H.-J. (2014) Natural single amino acid polymorphism (F19Y) in human galectin-8: detection of structural alterations and increased growth-regulatory activity on tumor cells. *FEBS Journal* **281**: 1446-1464.
- Rupert D. L. M., Claudio V., Lässer C. and Bally M. (2017) Methods for the physical characterization and quantification of extracellular vesicles in biological samples. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1861**: 3164-3179.
- Sahly H., Keisari Y., Crouch E., Sharon N. and Ofek I. (2008) Recognition of bacterial surface polysaccharides by lectins of the innate immune system and its contribution to defense against infection: The case of pulmonary pathogens. *Infection and Immunity* **76**: 1322-1332.
- Sauerbrey G. (1959) The use of quartz oscillators for weighing thin layers and for microweighing. *Zeitschrift für Physik* **155**: 206-222.

- Schirm M., Schoenhofen I. C., Logan S. M., Waldron K. C. and Thibault P. (2005) Identification of unusual bacterial glycosylation by tandem mass spectrometry analyses of intact proteins. *Analytical Chemistry* **77**: 7774-7782.
- Schmidt M. A., Riley L. W. and Benz I. (2003) Sweet new world: glycoproteins in bacterial pathogens. *Trends in Microbiology* **11**: 554-561.
- Schwartz-Albiez R. (2009) Inflammation and glycosciences. In *The Sugar Code. Fundamentals of Glycosciences*. H.-J. Gabius, Ed. Weinheim, Germany: Wiley-VCH. pp 447–466.
- Schwarz F. and Aebi M. (2011) Mechanisms and principles of N-linked protein glycosylation. *Current Opinion in Structural Biology* **21**: 576-582.
- Schweda E. K. H., Hegedus O. E., Borrelli S., Lindberg A. A., Weiser J. N., Maskell D. J. and Richard Moxon E. (1993) Structural studies of the saccharide part of the cell envelope lipopolysaccharide from Haemophilus influenzae strain AH1-3 (lic3 +). *Carbohydrate Research* **246**: 319-330.
- Schweda E. K. H., Richards J. C., Hood D. W. and Moxon E. R. (2007) Expression and structural diversity of the lipopolysaccharide of *Haemophilus influenzae*: Implication in virulence. *International Journal of Medical Microbiology* **297**: 297-306.
- Seaton B. A., Crouch E. C., McCormack F. X., Head J. F., Hartshorn K. L. and Mendelsohn R. (2010) Review: Structural determinants of pattern recognition by lung collectins. *Innate Immunity* **16**: 143-150.
- Senear D. F. and Teller D. C. (1981) Thermodynamics of concanavalin A dimer-tetramer self-association: sedimentation equilibrium studies. *Biochemistry* **20**: 3076-3083.
- Senfter D. and Mader R. M. (2015) Exosomes as novel biomarkers in anticancer therapy. *memo - Magazine of European Medical Oncology* **8**: 231-234.
- Seydel U., Koch M. H. J. and Brandenburg K. (1993) Structural Polymorphisms of Rough Mutant Lipopolysaccharides Rd to Ra from Salmonella minnesota. *Journal of Structural Biology* **110**: 232-243.
- Shafi R., Iyer S. P. N., Ellies L. G., O'Donnell N., Marek K. W., Chui D., Hart G. W. and Marth J. D. (2000) The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proceedings of the National Academy of Sciences* **97**: 5735-5739.
- Sharon N. (1996) Carbohydrate—lectin interactions in infectious disease. In *Toward anti-adhesion therapy for microbial diseases*. I. Kahane and I. Ofek, Eds. Boston, MA: Springer US. pp 1-8.
- Shenoy S. R., Barrientos L. G., Ratner D. M., O'Keefe B. R., Seeberger P. H., Gronenborn A. M. and Boyd M. R. (2002) Multisite and multivalent binding between cyanovirin-N and branched oligomannosides: Calorimetric and NMR characterization. *Chemistry & Biology* **9**: 1109-1118.
- Shibuya N., Goldstein I. J., Broekaert W. F., Nsimba-Lubaki M., Peeters B. and Peumans W. J. (1987) Fractionation of sialylated oligosaccharides, glycopeptides, and glycoproteins on immobilized elderberry (Sambucus nigra L.) bark lectin. *Archives of Biochemistry and Biophysics* **254**: 1-8.
- Shrive A. K., Tharia H. A., Strong P., Kishore U., Burns I., Rizkallah P. J., Reid K. B. M. and Greenhough T. J. (2003) High-resolution structural insights into ligand binding and immune cell recognition by human lung Surfactant Protein D. *Journal of Molecular Biology* **331**: 509-523.

- Simon M., Zieg J., Silverman M., Mandel G. and Doolittle R. (1980) Phase variation: evolution of a controlling element. *Science* **209**: 1370-1374.
- Simpson R. J., Kalra H. and Mathivanan S. (2012) ExoCarta as a resource for exosomal research. *Journal of Extracellular Vesicles* **1**: 10.3402/jev.v3401i3400.18374.
- Solís D., Bovin N. V., Davis A. P., Jiménez-Barbero J., Romero A., Roy R., Smetana Jr K. and Gabius H.-J. (2015) A guide into glycosciences: how chemistry, biochemistry and biology cooperate to crack the Sugar Code. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1850**: 186-235.
- Solís D. and Díaz-Mauriño T. (1997) Analysis of protein-carbohydrate interaction using engineered ligands. In *Glycosciences: status and perspectives*. H. J. Gabius and S. Gabius, Eds. Weinheim: Chapman & Hall. pp 345-354.
- Solís D., Fernández P., Díaz-Mauriño T., Jiménez-Barbero J. and Martín-Lomas M. (1993) Hydrogen-bonding pattern of methyl β-lactoside binding to the Ricinus communis lectins. *European Journal of Biochemistry* **214**: 677-683.
- Solís D., Romero A., Jiménez M., Díaz-Mauriño T. and Calvete J. J. (1998) Binding of mannose-6-phosphate and heparin by boar seminal plasma PSP-II, a member of the spermadhesin protein family. *FEBS Letters* **431**: 273-278.
- Solís D., Romero A., Menéndez M. and Jiménez-Barbero J. (2009) Protein–carbohydrate interactions: basic concepts and methods for analysis. In *The Sugar Code. Fundamentals of Glycosciences*. H.-J. Gabius, Ed. Weinheim, Germany: Wiley-VCH. pp 233-245.
- Song X., Lasanajak Y., Xia B., Smith D. F. and Cummings R. D. (2009) Fluorescent glycosylamides produced by microscale derivatization of free glycans for natural glycan microarrays. *ACS Chemical Biology* **4**: 741-750.
- Sperandio V., Torres A. G. and Kaper J. B. (2002) Quorum sensing *Escherichia coli* regulators B and C (QseBC): a novel two-component regulatory system involved in the regulation of flagella and motility by quorum sensing in *E. coli*. *Molecular Microbiology* **43**: 809-821.
- Stillmark H. (1888). Ueber Ricin, ein giftiges Ferment aus den Samen von *Ricinus comm*. L. und einigen anderen Euphorbiaceen, University of Dorpat.
- Stöber W., Fink A. and Bohn E. (1968) Controlled growth of monodisperse silica spheres in the micron size range. *Journal of Colloid and Interface Science* **26**: 62-69.
- Stowell S. R., Arthur C. M., Dias-Baruffi M., Rodrigues L. C., Gourdine J.-P., Heimburg-Molinaro J., Ju T., Molinaro R. J., Rivera-Marrero C., Xia B., Smith D. F. and Cummings R. D. (2010) Innate immune lectins kill bacteria expressing blood group antigen. *Nature Medicine* **16**: 295-301.
- Stowell S. R., Arthur C. M., McBride R., Berger O., Razi N., Heimburg-Molinaro J., Rodrigues L. C., Gourdine J.-P., Noll A. J., von Gunten S., Smith D. F., Knirel Y. A., Paulson J. C. and Cummings R. D. (2014) Microbial glycan microarrays define key features of host-microbial interactions. *Nature Chemical Biology* **10**: 470-476.
- Stowell S. R., Cho M., Feasley C. L., Arthur C. M., Song X., Colucci J. K., Karmakar S., Mehta P., Dias-Baruffi M., McEver R. P. and Cummings R. D. (2009) Ligand reduces galectin-1 sensitivity to oxidative inactivation by enhancing dimer formation. *Journal of Biological Chemistry* **284**: 4989-4999.

- Subra C., Laulagnier K., Perret B. and Record M. (2007) Exosome lipidomics unravels lipid sorting at the level of multivesicular bodies. *Biochimie* **89**: 205-212.
- Sumner J. B. (1919) The globulins of the jack bean, *Canavalia ensiformis*: preliminary paper. *Journal of Biological Chemistry* **37**: 137-142.
- Sumner J. B. and Howell S. F. (1936) Identification of hemagglutinin of jack bean with concanavalin A. *Journal of Bacteriology* **32**: 227-237.
- Swords W. E., Moore M. L., Godzicki L., Bukofzer G., Mitten M. J. and VonCannon J. (2004) Sialylation of lipooligosaccharides promotes biofilm formation by nontypeable *Haemophilus influenzae*. *Infection and Immunity* **72**: 106-113.
- Szathmary E. (2003) Why are there four letters in the genetic alphabet? *Nature Reviews Genetics* **4**: 995-1001.
- Takahashi H. (1992) The masking effect of sialic acid on Con A, PNA and SBA ectoderm binding sites during neurulation in the bantam chick embryo. *Anatomy and Embryology* **185**: 389-400.
- Tárnok A. (2008) SYTO dyes and histoproteins—myriad of applications. *Cytometry Part A* **73**: 477-479.
- Tchoupa A. K., Lichtenegger S., Reidl J. and Hauck C. R. (2015) Outer membrane protein P1 is the CEACAM-binding adhesin of *Haemophilus influenzae*. *Molecular Microbiology* **98**: 440-455.
- Tercero J. C. and Díaz-Mauriño T. (1988) Affinity chromatography of fibrinogen on *Lens culinaris* agglutinin immobilized on CNBr-activated Sepharose: study of the active groups involved in nonspecific adsorption. *Analytical Biochemistry* **174**: 128-136.
- Thery C., Ostrowski M. and Segura E. (2009) Membrane vesicles as conveyors of immune responses. *Nature Reviews Immunology* **9**: 581-593.
- Thiagarajan P. and Tait J. F. (1990) Binding of annexin V/placental anticoagulant protein I to platelets. Evidence for phosphatidylserine exposure in the procoagulant response of activated platelets. *Journal of Biological Chemistry* **265**: 17420-17423.
- Tombelli S. (2012) Piezoelectric biosensors for medical applications. In *Biosensors for Medical Applications*. Woodhead Publishing. pp 41-64.
- Töpfer-Petersen E., Romero A., Varela P. F., Ekhlasi-Hundrieser M., Dostàlovà Z., Sanz L. and Calvete J. J. (1998) Spermadhesins: a new protein family. Facts, hypotheses and perspectives. *Andrologia* **30**: 217-224.
- Trajkovic K., Hsu C., Chiantia S., Rajendran L., Wenzel D., Wieland F., Schwille P., Brügger B. and Simons M. (2008) Ceramide triggers budding of exosome vesicles into multivesicular endosomes. *Science* **319**: 1244-1247.
- Tullis R. H., Duffin R. P., Handley H. H., Sodhi P., Menon J., Joyce J. A. and Kher V. (2009) Reduction of hepatitis C virus using lectin affinity plasmapheresis in dialysis patients. *Blood Purification* **27**: 64-69.
- Turner A. P. F. (2000) Biosensors-Sense and Sensitivity. *Science* **290**: 1315-1317.
- Tzianabos A. O., Finberg R. W., Wang Y., Chan M., Onderdonk A. B., Jennings H. J. and Kasper D. L. (2000) T cells activated by zwitterionic molecules prevent abscesses induced by pathogenic bacteria. *Journal of Biological Chemistry* **275**: 6733-6740.

- Ulrich E. L., Akutsu H., Doreleijers J. F., Harano Y., Ioannidis Y. E., Lin J., Livny M., Mading S., Maziuk D., Miller Z., Nakatani E., Schulte C. F., Tolmie D. E., Kent Wenger R., Yao H. and Markley J. L. (2008) BioMagResBank. *Nucleic Acids Research* **36**: D402-D408.
- van Kooyk Y. and Rabinovich G. A. (2008) Protein-glycan interactions in the control of innate and adaptive immune responses. *Nature Immunology* **9**: 593-601.
- Van Valen L. (1974) Two modes of evolution. *Nature* **252**: 298-300.
- Varki A. (2006) Nothing in glycobiology makes sense, except in the light of evolution. *Cell* **126**: 841-845.
- Vasta G., Ahmed H., Nita-Lazar M., Banerjee A., Pasek M., Shridhar S., Guha P. and Fernández-Robledo J. (2012) Galectins as self/non-self recognition receptors in innate and adaptive immunity: an unresolved paradox. *Frontiers in Immunology* **3**.
- Vasta G. R. (2009) Roles of galectins in infection. *Nature Reviews Microbiology* **7**: 424-438.
- Velázquez-Campoy A., Ohtaka H., Nezami A., Muzammil S. and Freire E. (2004) Isothermal titration calorimetry. In *Current Protocols in Cell Biology*. John Wiley & Sons, Inc.
- Villalobo A. and Gabius H.-J. (1998) Signaling pathways for transduction of the initial message of the glycocode into cellular responses. *Cells Tissues Organs* **161**: 110-129.
- Vimr E. and Lichtensteiger C. (2002) To sialylate, or not to sialylate: that is the question. *Trends in Microbiology* **10**: 254-257.
- Virji M., Weiser J. N., Lindberg A. A. and Moxon E. R. (1990) Antigenic similarities in lipopolysaccharides of *Haemophilus* and *Neisseria* and expression of a digalactoside structure also present on human cells. *Microbial Pathogenesis* **9**: 441-450.
- Vretblad P. (1976) Purification of lectins by biospecific affinity chromatography. *Biochimica et Biophysica Acta (BBA) - Protein Structure* **434**: 169-176.
- Wang D., Liu S., Trummer B. J., Deng C. and Wang A. (2002) Carbohydrate microarrays for the recognition of cross-reactive molecular markers of microbes and host cells. *Nat Biotech* **20**: 275-281.
- Wang H., Head J., Kosma P., Brade H., Müller-Loennies S., Sheikh S., McDonald B., Smith K., Cafarella T., Seaton B. and Crouch E. (2008) Recognition of heptoses and the inner core of bacterial lipopolysaccharides by surfactant protein D. *Biochemistry* **47**: 710-720.
- Wang L., Cummings R. D., Smith D. F., Huflejt M., Campbell C. T., Gildersleeve J. C., Gerlach J. Q., Kilcoyne M., Joshi L., Serna S., Reichardt N.-C., Parera Pera N., Pieters R. J., Eng W. and Mahal L. K. (2014) Cross-platform comparison of glycan microarray formats. *Glycobiology* **24**: 507-517.
- Wang X., Matei E., Deng L., Koharudin L., Gronenborn A. M., Ramström O. and Yan M. (2013) Sensing lectin-glycan interactions using lectin super-microarrays and glycans labeled with dye-doped silica nanoparticles. *Biosensors and Bioelectronics* **47**: 258-264.
- Wang X., Norberg O., Deng L., Ramström O. and Yan M. (2011a) Synthesis of glyconanomaterials via photo-initiated coupling chemistry. In *Petite and Sweet:*

Glyco-Nanotechnology as a Bridge to New Medicines. American Chemical Society. pp 49-67.

- Wang X., Ramstrom O. and Yan M. (2011b) Dye-doped silica nanoparticles as efficient labels for glycans. *Chemical Communications* **47**: 4261-4263.
- Wang X., Ramstrom O. and Yan M. (2011c) Dynamic light scattering as an efficient tool to study glyconanoparticle-lectin interactions. *Analyst* **136**: 4174-4178.
- Wang X., Ramström O. and Yan M. (2010) Glyconanomaterials: synthesis, characterization, and ligand presentation. *Advanced Materials* **22**: 1946-1953.
- Waters P., Vaid M., Kishore U. and Madan T. (2009) Lung surfactant proteins A and D as pattern recognition proteins. In *Target pattern recognition in innate immunity*. U. Kishore, Ed.: Springer New York. pp 74-97.
- Watson J. D. and Crick F. (1953) Genetical implications of the structure of deoxyribonucleic acid. *Nature* **171**: 964-967.
- Weber T., Chandrasekaran V., Stamer I., Thygesen M. B., Terfort A. and Lindhorst T. K. (2014) Switching of bacterial adhesion to a glycosylated surface by reversible reorientation of the carbohydrate ligand. *Angewandte Chemie International Edition* **53**: 14583-14586.
- Weis W. I. and Drickamer K. (1996) Structural basis of lectin-carbohydrate recognition. *Annual Review of Biochemistry* **65**: 441-473.
- Weiser J. N., Love J. M. and Moxon E. R. (1989) The molecular mechanism of phase variation of *H. influenzae* lipopolysaccharide. *Cell* **59**: 657-665.
- Weiser J. N. and Pan N. (1998) Adaptation of *Haemophilus influenzae* to acquired and innate humoral immunity based on phase variation of lipopolysaccharide. *Molecular Microbiology* **30**: 767-775.
- Wight T. N., Heinegård D. K. and Hascall V. C. (1991) Proteoglycans: structure and function. In *Cell biology of the extracellular matrix*. New York: Springer Science+Business Media. pp 45-72.
- Williams B. A., Chervenak M. C. and Toone E. J. (1992) Energetics of lectin-carbohydrate binding. A microcalorimetric investigation of concanavalin A-oligomannoside complexation. *Journal of Biological Chemistry* **267**: 22907-22911.
- Winter L. E. and Barenkamp S. J. (2003) Human antibodies specific for the High-Molecular-Weight adhesion proteins of nontypeable *Haemophilus influenzae* mediate opsonophagocytic activity. *Infection and Immunity* **71**: 6884-6891.
- Wiseman T., Williston S., Brandts J. F. and Lin L.-N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry* **179**: 131-137.
- Woods R. J. (1995) Three-dimensional structures of oligosaccharides. *Current Opinion in Structural Biology* **5**: 591-598.
- Wright J. R. Pulmonary surfactant: a front line of lung host defense. *The Journal of Clinical Investigation* **111**: 1453-1455.
- Wu J. H., Herp A. and Wu A. M. (1993) Defining carbohydrate specificity of Ricinus communis agglutinin as Ga1β 1→ 4GlcNAc (II) > Galβ1→ 3GlcNAc (I) > Galα1→ 3Gal (B) > Galβ1→3GalNAc (T). *Molecular Immunology* **30**: 333-339.
- Xie B., Zhou G., Chan S.-Y., Shapiro E., Kong X.-P., Wu X.-R., Sun T.-T. and Costello C. E. (2006) Distinct glycan structures of uroplakins Ia and Ib:

Structural basis for the selective binding of FimH adhesin to uroplakin Ia. *Journal of Biological Chemistry* **281**: 14644-14653.

- Yan M. and Ren J. (2004) Covalent immobilization of ultrathin polymer films by thermal activation of perfluorophenyl azide. *Chemistry of Materials* **16**: 1627-1632.
- Yin Z. and Huang X. (2015) Boosting humoral immune responses to tumor-associated carbohydrate antigens with virus-like particles. In *Carbohydrates in Drug Design and Discovery*. S. Martin-Santamaria, J. Jimenez-Barbero and F. J. Canada, Eds.: The Royal Society of Chemistry. pp 132-150.
- Yoshioka Y., Konishi Y., Kosaka N., Katsuda T., Kato T. and Ochiya T. (2013) Comparative marker analysis of extracellular vesicles in different human cancer types. *Journal of Extracellular Vesicles* **2**: 20424.
- Zeng Z.-H., Castaño A. R., Segelke B. W., Stura E. A., Peterson P. A. and Wilson I. A. (1997) Crystal structure of mouse CD1: An MHC-like fold with a large hydrophobic binding groove. *Science* **277**: 339-345.
- Zhou G., Mo W.-J., Sebbel P., Min G., Neubert T. A., Glockshuber R., Wu X.-R., Sun T.-T. and Kong X.-P. (2001) Uroplakin Ia is the urothelial receptor for uropathogenic *Escherichia coli:* evidence from in vitro FimH binding. *Journal of Cell Science* **114**: 4095-4103.
- Zhou J., Hao N., De Zoyza T., Yan M. and Ramstrom O. (2015) Lectin-gated, mesoporous, photofunctionalized glyconanoparticles for glutathione-responsive drug delivery. *Chemical Communications* **51**: 9833-9836.
- Zonneveld M. I., Brisson A. R., van Herwijnen M. J. C., Tan S., van de Lest C. H. A., Redegeld F. A., Garssen J., Wauben M. H. M. and Nolte-'t Hoen E. N. M. (2014) Recovery of extracellular vesicles from human breast milk is influenced by sample collection and vesicle isolation procedures. *Journal of Extracellular Vesicles* **3**: 24215.
- Zuber C. and Roth J. (2009) N-glycosylation. In *The Sugar Code. Fundamentals of Glycosciences.* H. J. Gabius, Ed. Weinheim, Germany: Wiley-VCH. pp 87-109.

ANNEXES

Annex I

FT-IR spectra for SNP preparations

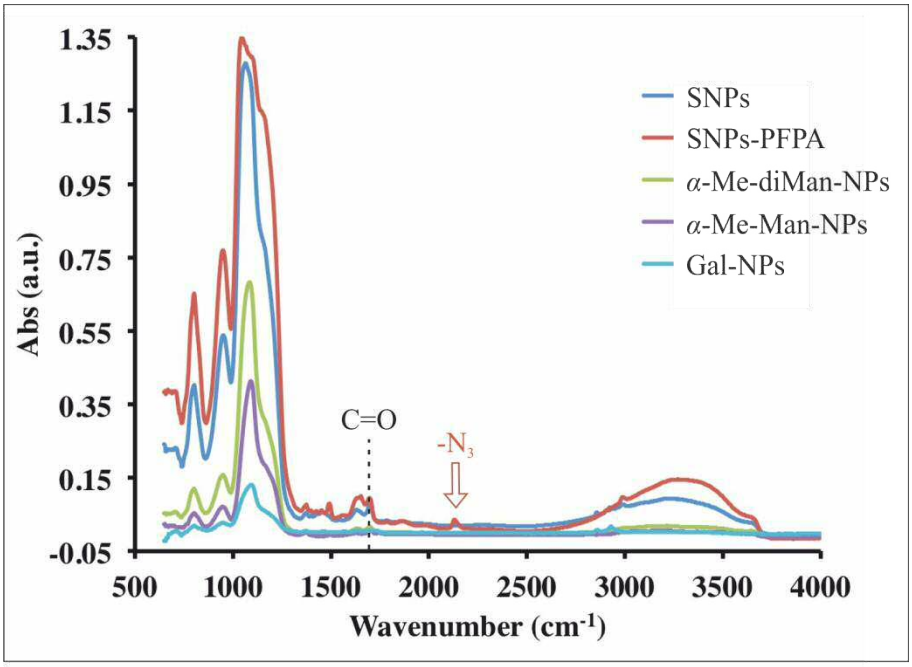


Fig. A1. The azide absorption signal at 2134 cm^{-1} was clearly observed on the SNPs-PFPA, as well as the amide C═O absorption band at 1639 cm^{-1} indicating successful functionalization of the SNPs with PFPA-silane. In the FT-IR spectrum of the final products the disappearance of the azide signal is characteristic as well as the appearance of triazole bands at 1697 and 1377 cm^{-1} for the sugar-derivatised GNPs (not shown due to spectra overlapping).

Annex II

Supplementary NMR spectra

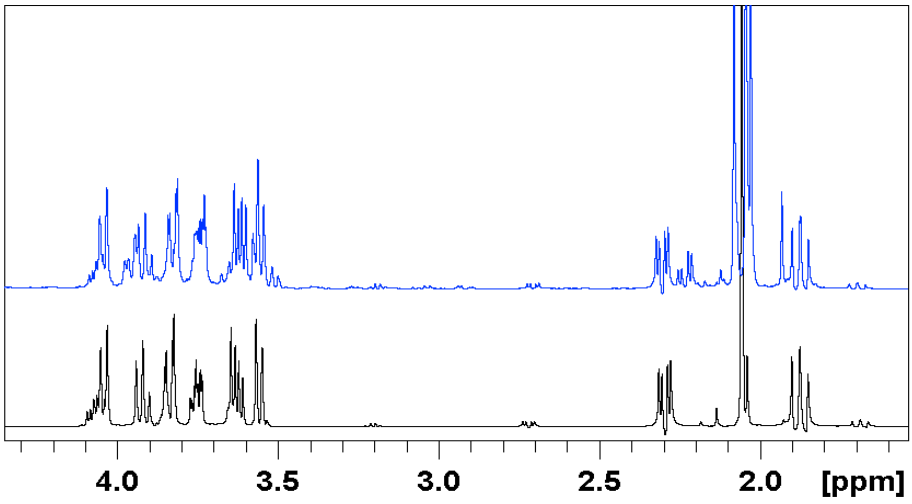


Fig. A2. ^{1}H-spectra of 10 mM Neu5Ac, before (black) and after acid hydrolysis (blue), performed as described in section 3.8.1. Spectra were acquired in deuterated PBS in a Bruker Avance 500 MHz spectrometer at 310 K, applying a water suppression pulse sequence with gradients (*zgesgp*), then referenced to the residual water (HDO) signal.

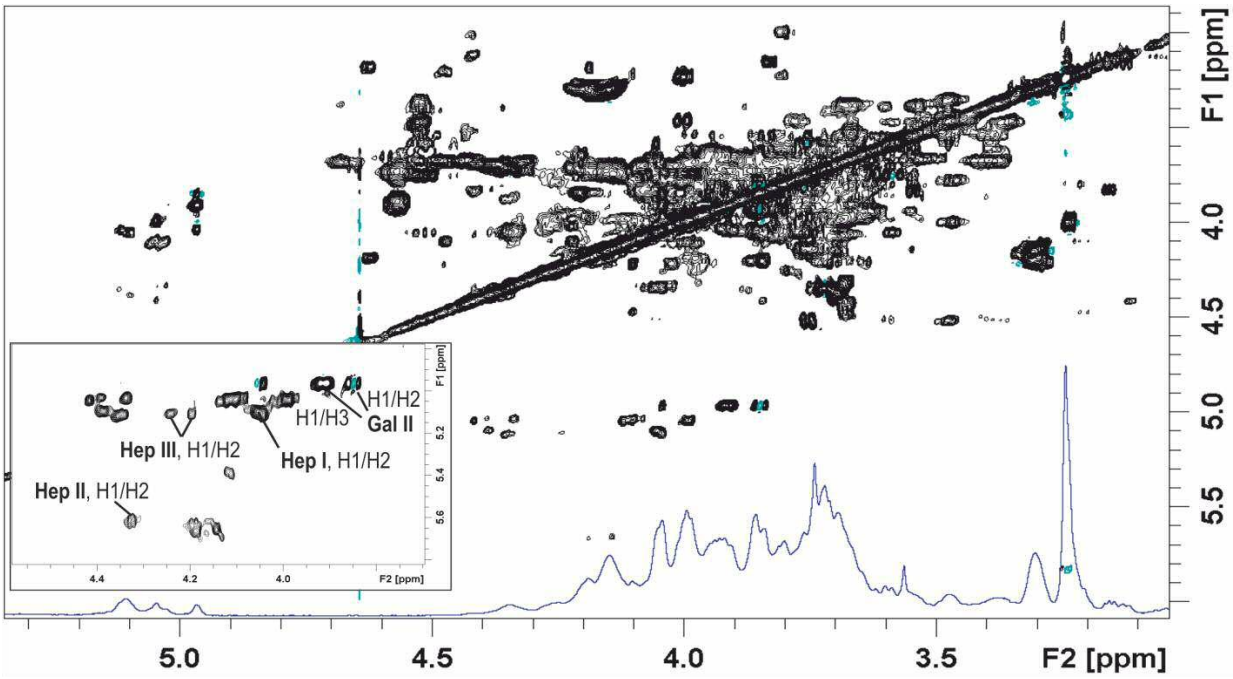


Fig A3. Section of the TOCSY spectrum of core oligosaccharides derived from Hex3 and Hex4 OS glycoforms of NTHi375Δ*ompP5*. Mixing time was of 70 ms. The corresponding 1D ^{1}H-spectrum was acquired applying a water suppression pulse sequence with gradients (*zgesgp*). Relevant correlation peaks are denoted in the zoomed section. All spectra were acquired at 310 K, in deuterated PBS in a Bruker Avance 600 MHz spectrometer equipped with a cryoprobe, then referenced to the residual water (HDO) signal, as described in the Materials and Methods (section 3.8).

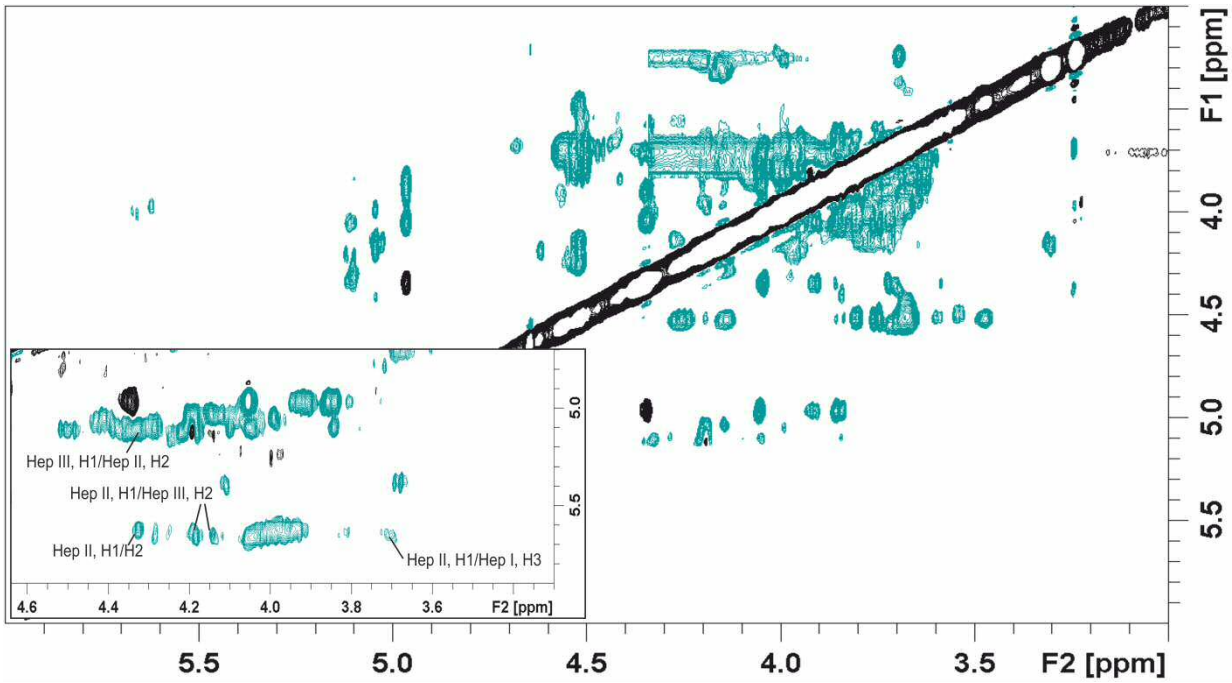


Fig A4. **Section of a ROESY spectrum of core oligosaccharides of NTHi375Δ*ompP5*.** Spectrum was acquired with a 300 ms spin-lock mixing time. In the inserted zoomed section of increased intensity, relevant intra- and inter-residual ROE contacts are denoted. Transglycosidic ROE cross-peaks between the proton pairs Hep III, H1/Hep II, H2 and Hep II, H1/Hep I, H3, established the sequence of the inner-core trisaccharide to be L-*α*-D-Hep*p*-(1,2)-L-*α*-D-Hep*p*-(1,3)-L-*α*-D-Hep*p*-(1→).

Table A1. ^{1}H-^{13}C NMR chemical shifts for core oligosaccharides derived from Hex3, Hex4 and Hex5 OS glycoforms of NTHi Rd

Isoform	Residue	Sugar unit	$^{3}J_{H,H}$	H1 (C1)	H2 (C2)	H3 (C3)	H4 (C4)	H5 (C5)	H6A (C6)	H6B	H7A (C7)	H7B
all	Hep I	3,4)-l-*α*-D-Hep*p*-(1→	N/R	5.09-5.14 (99.7)	4.04 (69.1)	N/D	4.24* (78.1)	N/D	4.14* (64.5)	N/D	N/D	N/D
all	Hep II	2)-l-*α*-D-Hep*p*-(1→	N/R	5.60-5.68 (99.2)	4.32 (78.8)	4.01 (87.8)	3.68 (72.8)	N/D	N/D	N/D	3.75 (65.0)	3.92 (61.2)
all	Hep III	2)-l-*α*-D-Hep*p*-(1→	N/R	5.10-5.13 (99.7)	4.24 (78.1)	4.02 (84.3)	3.80 (75.4)	N/D	N/D	N/D	N/D	N/D
all	Glc I	*β*-D-Glc*p*-(1→	8.2	4.52 (103.3)	3.38 (73.8)	3.48 (76.2)	3.55 (74.8)	3.61* (74.8)	4.14 (64.5)	4.27 (64.5)		
all	Glc II	4)-*β*-D-Glc*p*-(1→	7.5	4.67 (101.6)	3.38 (72.6)	3.69 (74.3)	3.69 (78.8)	3.69 (74.6)	3.85 (60.2)	3.99 (60.2)		
Hex3	Gal I	*β*-D-Gal*p*-(1→	7.7	4.47 (103.1)	3.55 (73.7)	3.68 (72.7)	3.94 (68.7)	3.74 (75.5)	3.78 (61.2)	N/D		
Hex4	Gal I	4)-*β*-D-Gal*p*-(1→	7.5	4.52 (103.3)	3.60 (71.1)	3.76 (72.3)	4.05 (77.6)	3.79* (75.5)	3.86 (60.5)	3.92 (60.5)		
Hex4	Gal II	*α*-D-Gal*p*-(1→	3.5	4.96 (100.4)	3.86 (68.7)	3.90 N/D	4.04 (69.1)	4.35* (71.0)	3.71-2 (60.7)			
Hex5	Gal II	3)-l-*α*-D-Gal*p*-(1→	3.7	4.94 (100.6)	3.90 N/D	3.96 (78.9)	4.26 (69.0)	N/D	N/D			
Hex5	GalNAc	*β*-D-GalNAc*p*-(1→	N/R	4.64 (100.2)	3.95 (68.7)	3.90 (69.3)	4.04 (69.2)	3.69* (71.0)				
all	PEtn			4.16 (62.1)	3.22 (40.8)							
all	PCho			4.38 (59.6)	3.69 (66.2)							

$^{3}J_{H,H}$ values for anomeric ^{1}H resonances (H1) are given. Except where indicated, chemical shifts are the same (± 0.02 ppm) for the conserved inner core residues, shared by both glycoforms. N/R: not resolved, N/D: not defined.

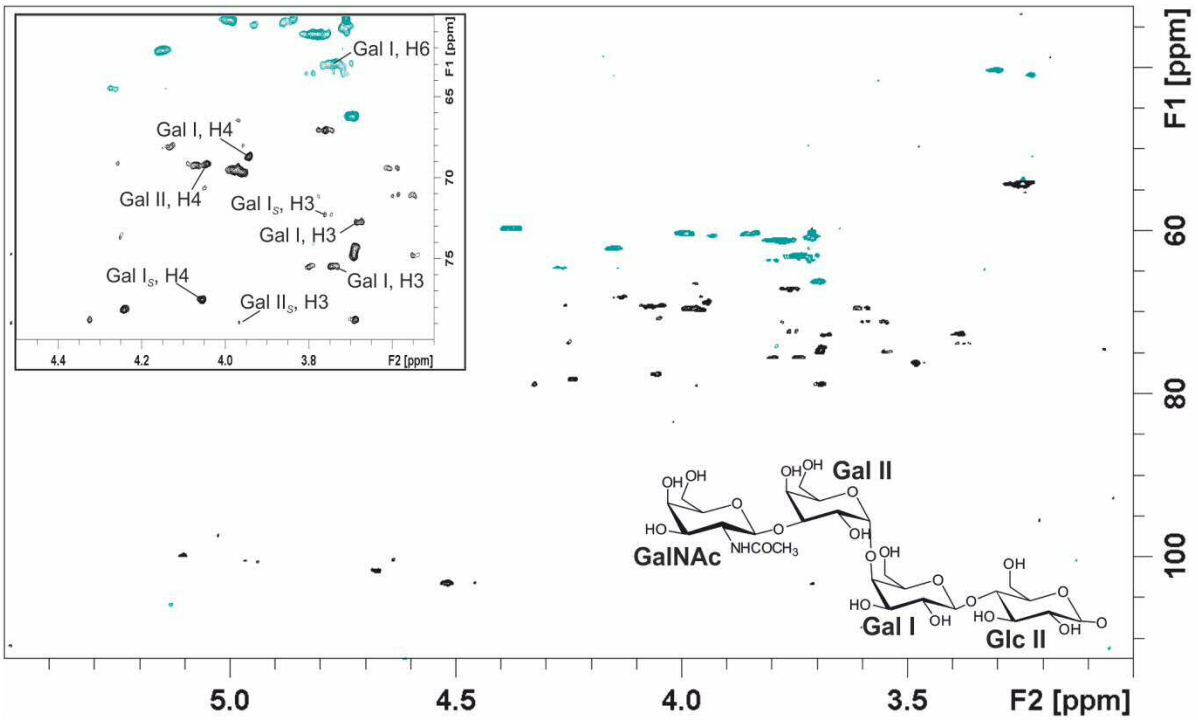


Fig. A5. Section of the multiplicity edited ^{1}H-^{13}C HSQC spectrum of core oligosaccharides derived from a mixture of OS glycoforms of the Rd strain. The inserted spectral zoom at δ ~ 3.5-4.5 ppm (^{1}H, F2 axis) corresponds to a zone of special interest for galactose moieties, where several resonances were successfully assigned, and here denoted for relevant protons. Symbol *S* is used as for substituted glycans, given the coexistence of multiple glycoforms in the sample. Chemical shifts and coupling constants of the anomeric protons are reported in Table A1. For both spectra: CH_2 are shown in petrol; CH and CH_3 in black.

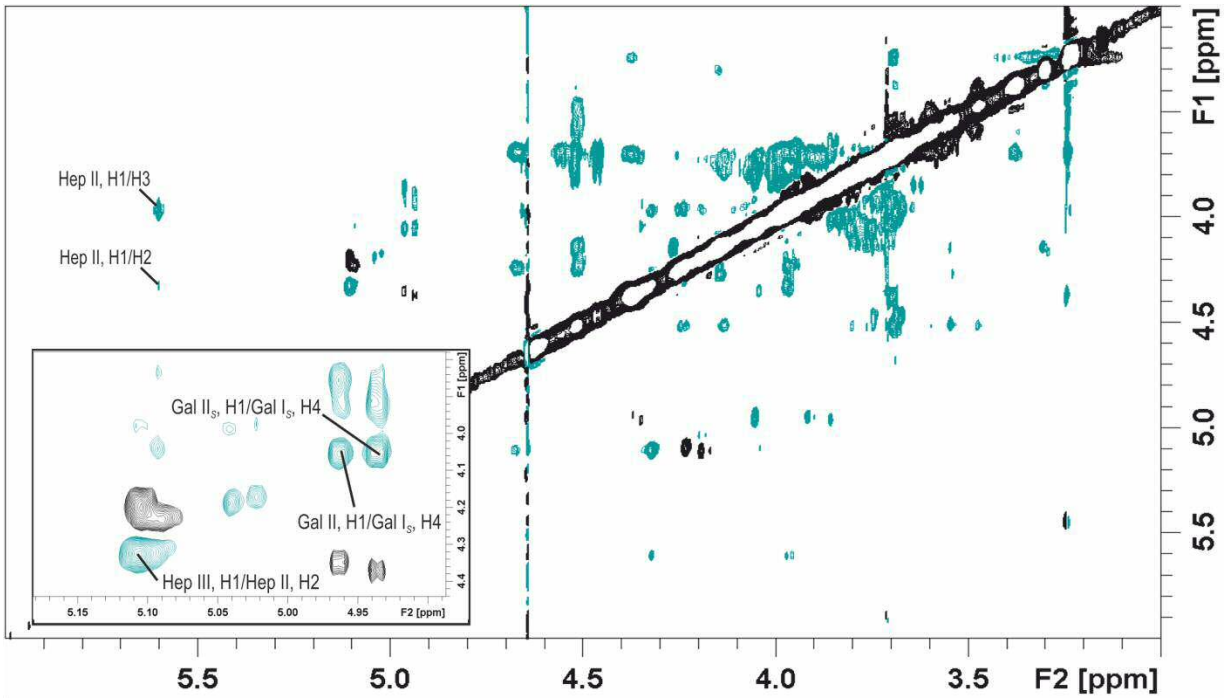
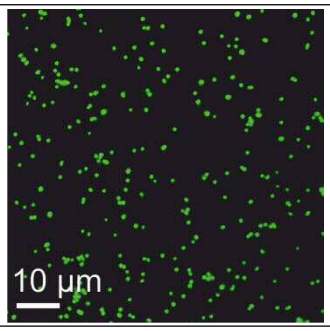
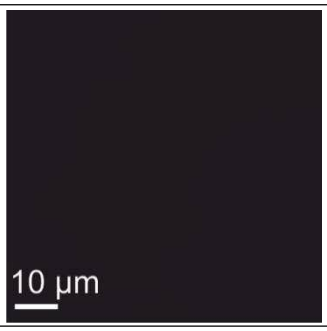
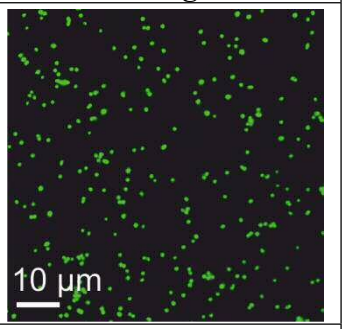
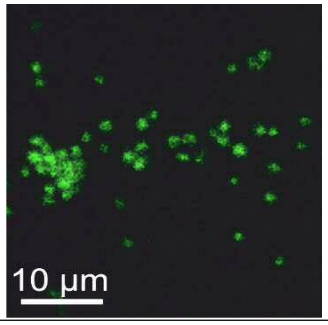
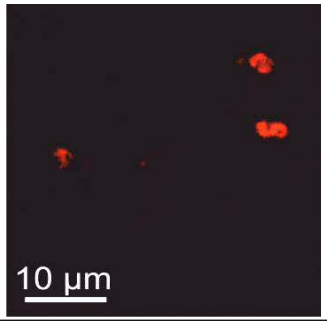
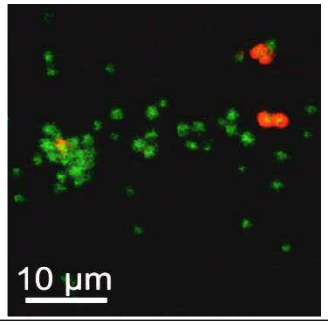
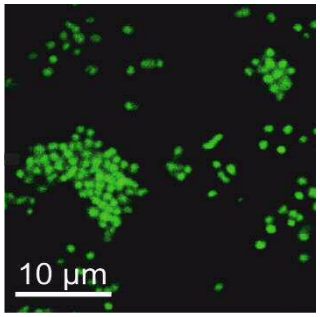
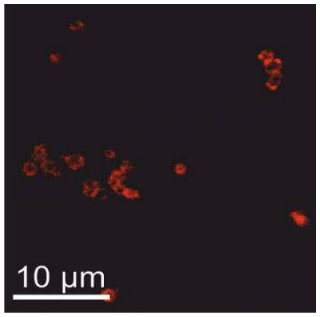
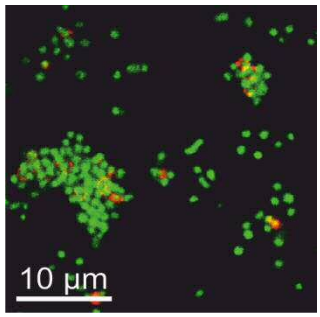
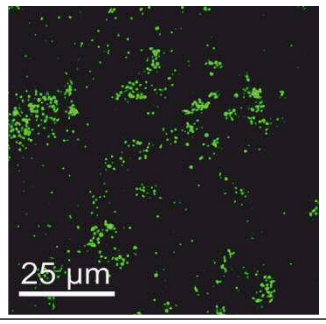
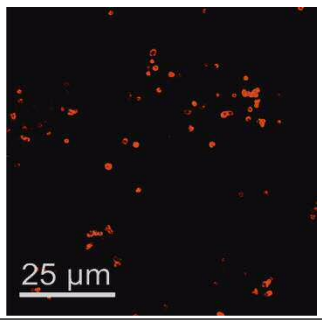
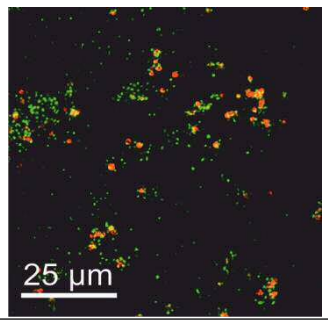
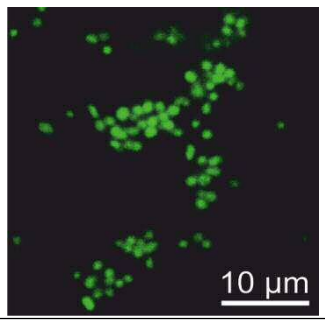
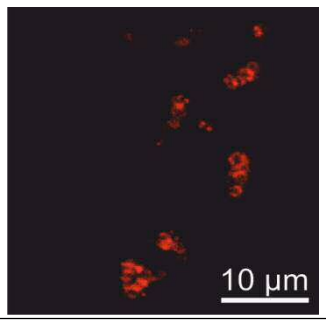
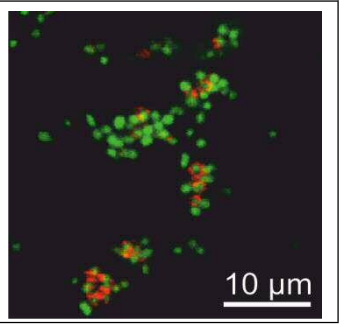


Fig A6. Section of a ROESY spectrum of core oligosaccharides of the Rd strain.
Spectrum was acquired with a 300 ms spin-lock mixing time. Relevant intra- and inter-residual ROE contacts are denoted. In the inserted zoomed section, symbol S stands for substituted residues. Substitution of the Gal II residue in position 3, provokes a small, yet perceptible shift for its H3 resonance.

Annex III

Table A2. Confocal microscopy photos for different galectins (red) and NTHi375 (green)

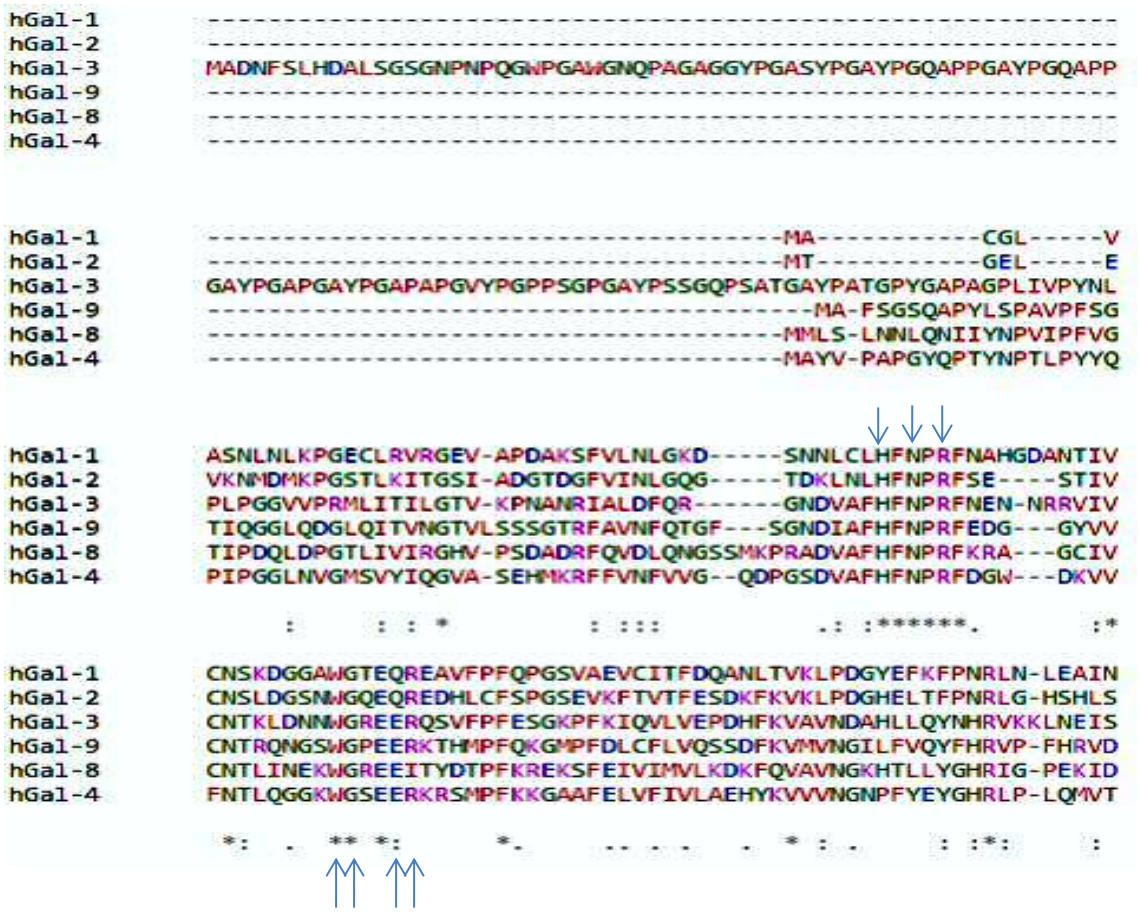
	Excitation at 488 nm	Excitation at 561 nm	Merge
hGal-1			
hGal-3			
hGal-4			
hGal-4*			
hGal-8			

*Bacteria were incubated with hGal-4 in absence of dithiothreitol.

Annex IV

Sequence alignments for galectins used in this Thesis

As briefly described in the introduction of the section 4.10, human galectins were classified into three groups on the basis of their structure: proto-type, chimeric, and tandem-repeat. A galectin-CRD bears approximately 130 amino acids. Sequence alignments for hGal-1, -3, -4, -8 and -9 here shown, were acquired using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). For tandem-repeat galectins (-4, -8 and -9) several isoforms have been identified, bearing distinct length peptide-linker, which is prone to peptidase cleavage. Residues that are highly conserved between galectins and those that are known to make contacts with carbohydrate ligands are indicated by arrows. Binding domains for tandem-repeat galectins are in boxes.



Annex V

Calculation of the dissociation constants for GNPs

Calculation of the dissociation constants for the ConA-GNPs (α-Me-Man- and α-Me-diMan-NPs) interaction took place as follows, based on a comparable set-up previously established (Rivera-Sagredo *et al.*, 1991).

Here, [A] corresponds to MAF285 (biotinylated heptylmannoside) and [B] to the GNPs concentration, K_A to the dissociation constant for MAF285 and K_B to that of GNPs. Assuming that the association constant, K_a, of MAF285 for ConA is comparable to that measured for α-Me-Man by ITC, $(8.77 \pm 0.09) \times 10^3$ (Table 4.13), it corresponds to $1/K_A$. Thus, $K_A = 1.14 \times 10^{-4}$.

If M_{total} stands for the total amount of ConA in the incubation mixture, M_o the amount of free ConA, and M_A, M_B the amount of ConA-sites captured by A and B respectively, then $M_{total} = [M_o] + [M_A] + [M_B]$ (eq. 1).

Eq. 1 can be substituted in terms of the dissociation constants K_A and K_B of the complexes of ConA with A and B respectively, referred to as [MA] and [MB]:

$$K_A = \frac{[\mathrm{M}]\times[\mathrm{A}]}{[\mathrm{MA}]} \qquad K_B = \frac{[\mathrm{M}]\times[\mathrm{B}]}{[\mathrm{MB}]}$$

rearranged to:

$$[\mathrm{M_o}] = \frac{K_A\times[\mathrm{MA}]_\mathrm{o}}{[\mathrm{A}]_\mathrm{o}} \qquad [\mathrm{MB}] = \frac{[\mathrm{M}]_\mathrm{o}\times[\mathrm{B}]}{K_B}$$

For the binding experiment of A to ConA we have:

$$\mathrm{M_{total}} = \mathrm{M_o} + [\mathrm{MA}]_\mathrm{o} = \frac{K_A\times[\mathrm{MA}]_\mathrm{o}}{[\mathrm{A}]_\mathrm{o}} + [\mathrm{MA}]_\mathrm{o}\left(= [\mathrm{MA}]_\mathrm{o}\right)\ 1+\frac{K_A}{[\mathrm{A}]}$$

$$[\mathrm{MA}]_\mathrm{o} = \mathrm{M_{total}} \Big/ 1+\frac{K_A}{[\mathrm{A}]}$$

Thus, eq. 1 is rearranged to:

$$\mathrm{M_{total}} = \frac{K_A\times[\mathrm{MA}]}{[\mathrm{A}]} + [\mathrm{MA}] + \frac{[\mathrm{M}]_\mathrm{o}\times[\mathrm{B}]}{K_B}$$

$$\mathrm{M_{total}} = \frac{K_A\times[\mathrm{MA}]}{[\mathrm{A}]} + [\mathrm{MA}] + \frac{[\mathrm{B}]}{K_B}\times\frac{K_A\times[\mathrm{MA}]}{[\mathrm{A}]}$$

$$\mathrm{M_{total}} = [\mathrm{MA}]\left(1+\ \frac{K_A}{[\mathrm{A}]}\left(1+\ \frac{[\mathrm{B}]}{K_B}\right)\right)$$

$$[\mathrm{MA}] = \mathrm{M_{total}} \Big/ 1+\left(\frac{K_A}{[\mathrm{A}]}\left(1+\ \frac{[\mathrm{B}]}{K_B}\right)\right)$$

$$\frac{[\mathrm{MA}]}{[\mathrm{MA}]_\mathrm{o}} = 50\,\% = \frac{\mathrm{M_{total}} \Big/ 1+\left(\frac{K_A}{[\mathrm{A}]}\left(1+\frac{[\mathrm{B}]_{50\%}}{K_B}\right)\right)}{\mathrm{M_{total}} \Big/ 1+\frac{K_A}{[\mathrm{A}]}} = \frac{1+\frac{K_A}{[\mathrm{A}]}}{1+\left(\frac{K_A}{[\mathrm{A}]}\left(1+\frac{[\mathrm{B}]_{50\%}}{K_B}\right)\right)}$$

$$0.5 + 0.5 \times \frac{K_A}{[A]} + 0.5 \times \frac{K_A}{[A]} \times \frac{[B]_{50\%}}{K_B} = 1 + \frac{K_A}{[A]}$$

$$0.5 \times \frac{K_A}{[A]} \times \frac{[B]_{50\%}}{K_B} = 0.5 + 0.5 \times \frac{K_A}{[A]}$$

$$\frac{K_A}{[A]} \times \frac{[B]_{50\%}}{K_B} = 1 + \frac{K_A}{[A]}$$

$$\frac{[B]_{50\%}}{K_B} = 1 + \frac{[A]}{K_A}$$

$$K_B = \frac{[B]_{50\%}}{1 + \frac{[A]}{K_A}}$$

The concentration of free ligands [A] and [B] can be considered equal to the total amount present in the incubation mixture, since their concentrations are, in all cases, much larger than the concentration of ConA. Therefore, a plot of the reciprocal of the ConA fraction bound to A versus the B should give a straight line since [B] is a fixed value, being K_B = (intercept-l)/slope.

Annex VI

Localization of N-glycosylated Asn residues in three HMW1 adhesins of NTHi

The Asn residues of HMW1 of strain 12 are shown in comparison with those from 86028-NP and NTHi375. Conservative and non-conservative amino acid substitutions are shown in a grey and black background respectively (Domenech *et al.*, 2016).

Strain 12 (AAA20527)	Strain 86-028NP (AAX88733)	Strain Hi375 (AIT67726)
N636	N630	I640
N642	S636	H646
N709	N704	T715
N806	F801	F808
N912	N906	Q915
N928	I922	K931
N946	S940	D949
N952	H946	H955
N964	N958	D967

Multiple partial alignment of the HMW1 adhesins of strains 12 (AAA20527), 86-028NP (AAX88733) and 375 (AIT67726). The locations of N-glycosylated Asn residues were previously reported (Gross *et al.*, 2008), and here the conserved (black triangles) and not conserved ones (red triangles) are highlighted for the three aligned proteins (Domenech *et al.*, 2016).

.