ASSAY DEVELOPMENT AND KINETIC STUDIES OF PROTEIN-SMALL MOLECULE INTERACTIONS

By

BRIAN EMBER

(Under the direction of Geert-Jan Boons)

ABSTRACT

Rational drug design remains one of the best methods in the development of new pharmaceutical compounds. The cyclic research pattern of testing and derivatizing these compounds requires the ability to rigorously study the biomolecular kinetics of both the natural substrate – protein interaction as well as that of the derivatives with the native protein. Regardless of whether the desired outcome is inhibition, signal transduction or enhanced avidity, greater knowledge of the system will facilitate better medicinal compounds in a shorter period of time.

Over the past few decades, the importance of oligosaccharides as important biological ligands in a variety of immune associated and disease related pathways have made them attractive targets for synthetic medicinal chemists. Along with the ability to alter selectivities via changes in linking patterns and derivitization, utilization of multivalency, found throughout nature with sugar interactions, greatly enhances affinity without the need of a tight-binding ligand. The increased affinities of multivalent saccharides have been exploited in the synthetic design of inhibitors for bacterial toxins as well as blocking the binding for viruses to cells.

This dissertation combines the efforts of studying small molecule – protein interactions with the importance of oligosaccharide chemistry as biological markers in synthesis as well as assay development. Monovalent and polyvalent synthetic substrates containing terminal sialic acid were investigated as to their effectiveness as a substrate toward three modular sialidases and one without a carbohydrate binding domain. The resulting novel mechanism unraveled resulted in the design of novel selective inhibitors for bacterial sialidases. Furthermore, new studies of human proteins believed to take part in immune recognition were assayed with surface plasmon resonance technology utilizing glycopeptide part structures synthesized in the Boons' laboratory. The kinetic studies of these part structures led to the understanding that one of the proteins studied is highly selective toward the recognition of DAP-type peptidoglycan.

Index Words: Carbohydrate, Multivalency, Sialidase, Neuraminidase, Kinetics, Peptidoglycan, Surface Plasmon Resonance, Peptidoglycan Recognition Proteins

ASSAY DEVELOPMENT AND KINETIC STUDIES OF PROTEIN-SMALL

MOLECULE INTERACTIONS

By

BRIAN EMBER

B.S., The University of Florida, 1998

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial

Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2006

DEDICATION

I have enjoyed myself, many say too much, during my years of higher education. One more football season turned into twelve total, including numerous SEC championships and one national title. It is to the people that supported me most during those many years, those who know me not as a scientist but as a person, that I dedicate this dissertation.

I will begin with my friends, the people I have (kind of) grown up with becoming the man I am today. (Blame everything on them.) After all, without fun, what is life? Tony and Rick, although I will never truly be a Gator, the times we had (and continue to have) are some of the best of my life. And no matter how successful you guys become, I will remember when you were 18 year old punks. More importantly, I will make you call me doctor forever.

Johnboy and Charles, it was with you that I realized what a nerd I really am. Drinking and talking about chemistry...it would be embarassing to almost anyone else. Paul, I can honestly say that I would not have remained sane without you. You were there during the best of times and the worst and I think you can say the same of me. We studied together, we partied together and we became adults together. Thanks. And finish!

I consider myself extremely lucky, for I have more grandparents that love me than most people. I dedicate this dissertation to each of them. To my Grandmother Mary for always being right around the corner when I needed her. To my Nana for teaching me to take time to enjoy life (and to introducing me to the state of Florida). To my Grandfather

Ed for introducing me to my one true addiction, Georgia football. Andrea may curse you for it, but as an addiction, it could be worse. And to Connie for teaching me to shoot spit balls with the best of them.

Stacie and Brandon, you understand my last few years better than anyone else in the family as you lived something very similar. You also made Athens home, for I cannot imagine how difficult my transition there would have been without the two of you. Stacie, we get along wonderfully (now), and although I know that it does not need to be said, I want to tell you that I am proud of you and that I love you.

Mom and Dad, I dedicate this dissertation to you. There are so many things that I could say, but nothing would capture the appreciation I have for the way you have raised, supported and loved me. I guess the best way to put it into words is to say that I hope someday that Andrea and I can be half the parents that you were to Stacie and I. And Go Dawgs!

And to Dre, my best friend and the love of my life, I dedicate this work. You are the only person who truly knows what it took to obtain this degree. Many were there through the good times and the successes, but no one had to endure the stress and the sleepless nights except you and I. By focusing so hard on finishing, I have left you to care for the house, for Staley and to plan an overseas wedding while finding time to go to school and work. When I was not in the best of moods, you were there to take my attitude and you were there to settle me down. I am sorry. I love you with all of my heart. Regardless of what waits for us in the future, I will be happy because it will be with you. Andrea, I dedicate my life to you.

ACKNOWLEDGEMENTS

A countless number of mentors, colleagues, acquaintances, friends and family have assisted me along the lengthy journey of formal education ending (for now) with this dissertation. Although I will not be able to mention all of you, I would like to begin with a general thank you to everyone who has been there.

Specifically, I begin with my most recent scientific mentor and major professor Geert-Jan Boons. Thank you for your expertise, guidance and support as well as providing me with the chance to attain this degree by entrusting me with two exciting projects and a lab environment in which to learn. I would also like to take this opportunity to thank my committee members, professors Robert Woods and Robert Phillips, for taking the time out of their busy schedules to attend my seminars and assist with my orals, dissertation and defense as well as the many other ways you have helped during my time at the UGA.

I also want to personally name many of the others who have helped my scientific career to this point. From the University of Florida, I would like to thank Dr. Merle Battiste and Dr. Desmond Kwan as well as Julie Ann Steffens. From the University of Georgia Chemistry department, I would like to thank John Britton. Among the many friends I have made at the Complex Carbohydrate Research Center at UGA, I want to mention (doctors) Aloysius Siriwardena, Smita Thobhani, Martin McWatt, Parastoo Azadi, Abhijit Roychowdhury and Sanjay Kumar. I would also like to thank Trina Abney, Charles Stanton and Paul Dietzel. Thanks also to Dr. Eric Roush from Biacore. Finally, I would like to thank the University of Florida and the University of Georgia.

Table of Contents

Page

List of Abbreviations

Ac	Acetyl
Ac_2O	Acetic anhydride
Bn	Benzyl
dd	double of doublets
DCM	methylene chloride
DMF	dimethyl formamide
Et	ethyl
EtOH	ethanol
FAB-MS	fast atom bombardment mass spectroscopy
GlcNAc	*N*-Acetyl glucosamine
HPAEC	high pH anion exchange chromatography
HPLC	high performance liquid chromatography
h	hour
Hz	hertz
m	multiplet
MeOH	methanol
m/z	mass to charge ratio
Me	methyl
min	minute
mM	millimolar
mmol	millimole
MS	molecular sieves

NMR	nuclear magnetic resonance
PGN	peptidoglycan
ppm	parts per million
R_f	retention factor
s	singlet
t	triplet
TFA	trifluoroacetic acid
TLC	thin layer chromatography

Chapter 1

Introduction

Rational drug design remains one of the best methods in the development of new pharmaceutical compounds. The cyclic research pattern of testing and derivatizing these compounds requires the ability to rigorously study the biomolecular kinetics of both the natural substrate – protein interaction as well as that of the medicinal derivative with the native protein. Whether the desired effect is inhibition, signal transduction or enhanced avidity, greater knowledge of the system will facilitate better drugs in less time.

1.1 Carbohydrates

Carbohydrates are aldehyde or ketone compounds containing multiple hydroxyl groups. Once believed to be primarily a source of energy and structural component of living organisms, they have become well established as substrates that play key roles in many biological events. As chief constituents in a myriad of biological processes, carbohydrates have become a focus for an ever increasing number of research studies centering on their mechanistic biological roles as well as the synthesis of both naturally occurring and derivatized compounds (1). Among the known roles of carbohydrates in physiological events include involvement in neuronal development and cell proliferation (1). More specifically, the alterations in the carbohydrate makeup of cell surfaces that have been through oncogenic transformations appear to be intimately associated with the invasiveness and metastatic potential of malignant cells (2-5). Carbohydrates are also involved in the recognition and immunological response necessary for survival during pathogenic infection (6-8). Blood coagulation at a site of injury is known to require carbohydrate assistance via heparin, a sulfated oligosaccharide (9). Oligosaccharides have been associated with both the development of new organisms including fertilization (1) and embryogenesis (10-11) as well as everyday working functions such as hormone circulation (12) and enzyme routing (13).

The importance of glycobiology, including but not exclusive to the examples of carbohydrate mediated biological events listed above, has been rationalized by two rudimentary views (14). First, most cells are covered with carbohydrates making them very accessible to extracellular interactions. Secondly, due to the massive structural diversity designed into the carbohydrate makeup through the ability of glycosidic bonds to be formed via several external functional hydroxyls per monosaccharide unit plus anomeric diversity, carbohydrates have the capability of carrying vast quantities of biological information making them a highly desirable subset for biomolecular testing and drug design (14-15).

1.2 Carbohydrate – Protein Recognition and Interaction

Carbohydrates interact with proteins, including enzymes, immunoglobins and recognition proteins given the umbrella term lectins and immunoglobins, which initiate the immune response (14). Thus, the mechanistic study of carbohydrate recognition is of principal interest in the medicinal community for developing novel drugs that initiate or inhibit such processes. To understand these mechanisms, a basic knowledge of the structural interactions of sugars and proteins at an atomic level is needed.

1.2.1 Structure and Recognition

The core structural features factoring into oligosaccharide recognition by proteins are hydrogen bonding, metal complexation, ionic effects, van der Waals forces and hydrophobic interactions (16-18). Hydrogen bonds in carbohydrate complexes to lectins, enzymes and antibodies, according to x-ray crystallographic analyses, occur between the sugar hydroxyl groups and either the amide functionality of the protein residues, which act as hydrogen bond acceptors, or with acidic protein groups, which act as hydrogen bond donors. Co-operative hydrogen bonding occurs when a sugar interact as both an acceptor and donating group simultaneously (Figure 1.1). Protein residues with capable side chains including aspartic acid and arginine can also form bidentate hydrogen bonds with monosaccharides either with vicinal hydroxyl groups or with one hydroxyl and the

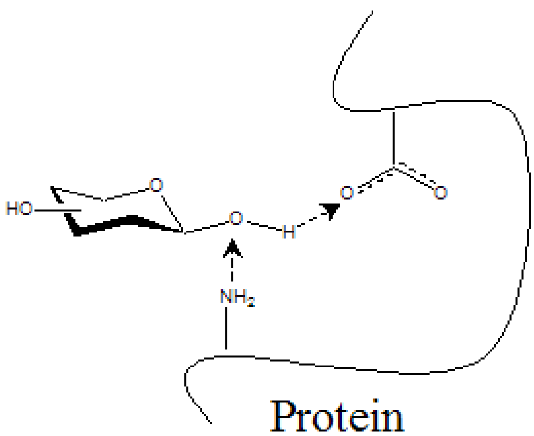


Figure 1.1: Co-operative hydrogen bonding between sugar and protein

ring oxygen (17) (Figure 1.2). Water molecules can also form bridges between a protein and sugar through hydrogen bonding, often with the resulting interaction being as stable as a direct bond. (19). Furthermore, water molecules can be found in both free lectin sites as well as bound sugar-protein complexes (20), thus acting as fixed protein extensions (17).

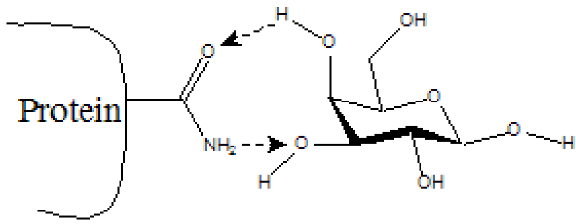


Figure 1.2: Bidentate hydrogen bonding between sugar and protein

Many carbohydrate interactions to both lectins and enzymes require metal cations for biological function (21-22). More specifically, a calcium ion coordinates to mannose hydroxyl groups concurrently with a residue of mannose binding protein A (MBP-A) forming a bridged ligation (18). Other proteins requiring ions for carbohydrate recongnition, such as the calcium dependence of all class 1 mannosidases (22), utilize coordination to organize the protein, the substrate or both into conformations proper for complexation (23).

Naturally charged sugars, such as sialic acid, along with sulfated and phosphorylated carbohydrates, can use ionic charges, both to attract and repel, to participate in carbohydrate-protein binding asscociations. For example, the negatively charged carboxylate of the sulfated polysaccharide heparin forms a salt bridge between the positively charged side chains of the lysine and arginine residues in antithrombin

(24). Sialic acids on cell surfaces utilize ionic charge to both repel unwanted pathogenic proteins while attracting required self proteins such as sialidases (25).

Hydrophobic interactions, often through the participation of aromatic amino acids, confer specific binding to protein binding sites by prohibiting non-substrate sugars from accessing the site via steric hindrances and unfavored non-polar interactions (26). These aromatics stack against the face of the sugar rings, thus allowing only the correct residues to form an association. For example, in the X-ray crystal structure of MBP-A, both glucose residues of maltose are sandwiched between aromatic rings (27). There is also stacking of the apolar side of the galactose ring against the plane of either tryptophan or phenylalanine rings in most galactose-protein complexes (17).

1.3 Multivalency

The majority of saccharide ligands bind with poor affinity to protein receptors (28). Dissociation constants (Kd) are often found in the millimolar to micromolar range, well above the generally accepted nanomolar affinity required for most biological recognition (14,29). In nature, the low affinity of carbohydrate-protein interactions is often overcome through multivalent interactions. Multivalency is the simultaneous interaction of multiple ligands with multiple receptors, increasing the affiity of, in this case carbohydrates, to a level necessary for biological recognition (Figure 1.3) (30). Proteins containing more than one binding site combined with carbohydrate ligands that are displayed as clusters on glycolipids or cell-surfaces or on polymers or other scaffolds can thus bind in a multivalent fashion. While an additive effect would slightly increase the affinity, often much less than the recognition requirements necessitate, multivalent binding increases the affinity well above the monomeric ligand-receptor interaction on a valence-corrected basis. Beyond inter- and intramolecular interactions, an “avidity entropy” described by Kitov and Bundle, gives a thermodynamic basis for the increase in affinity associated with multivalency (31). Increases in affinity have also been attributed to entropy contributions obtained by the expulsion of water molecules at the binding interface (32-33). The overall phenomenon was discovered by Lee et al in 1995 and is termed the *glycoside cluster effect* (14,34).

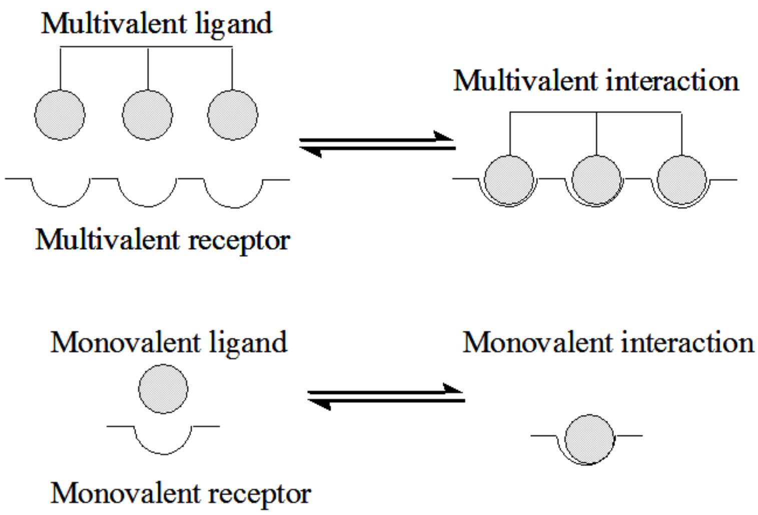


Figure 1.3: Monovalent versus multivalent interactions

1.4 Effects of Multivalency

The term multivalency, although always referring to multiple ligands on a single moiety binding with multiple receptors on a second entity, encases a number of related but different effects which all increase the affinity of the interaction when compared on a molar basis to the monovalent counterpart. The following section will discuss the different types of interactions in which enhancements in affinities of multivalent ligands takes place.

1.4.1 The Chelate Effect

The chelate effect describes the secondary (or tertiary, etc) occupation of a multivalent ligand to an oligomeric receptor after initial binding takes place due to a favorable orientation of the molecules. This interaction requires an initial intermolecular association followed by one or more intramolecular associations (Figure 1.4) (35). Accordingly, loss of translational entropy occurs during the initial binding event, but not for the subsequent intramolecular associations. However, there is a loss of conformational flexibility as increased intramolecular binding events occur that also affects free energy of binding.

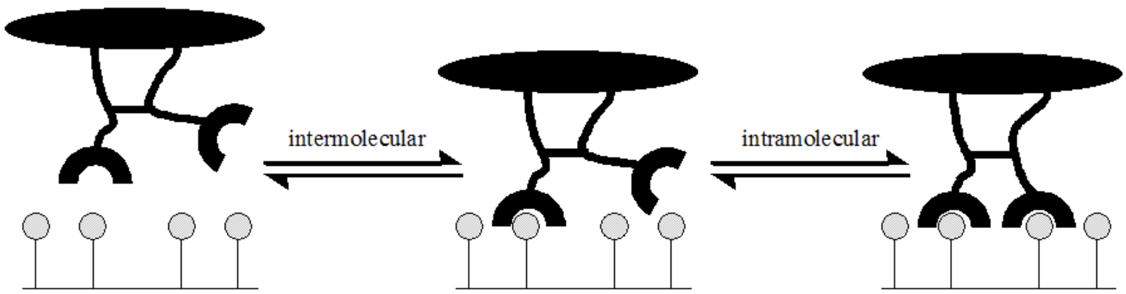


Figure 1.4: The chelate effect

1.4.2 The Statistical Effect

Although the theory of a local concentration, in which the area directly surrounding a multivalent display of ligands is recognized by receptors as an actual higher concentration of the ligand, is controversial, the sheer statistics of higher densities of ligands and receptors cannot be denied. The resulting increase in functional affinities, a term used to describe the measured activity of multivalent ligands compared to monovalent ligands as well as other multivalent compounds, usually as the apparent association (K_a) or inhibition constant (K_d), is theorized as the increase in the probability of rebinding versus dissociation (Figure 1.5). In terms of the equilibrium association constant K_a with:

$$K_a = k_{on} / k_{off} \qquad \{1.1\}$$

and k_{on} defined as the second-order rate constant for complex formation while k_{off} is the first-order rate constant for complex dissociation, a reduction in dissociation of the complex thus increases the functional affinity of the interaction. As a result, higher binding affinities exhibited by multivalent interactions is partially due to lower off-rates (36-37). Unlike the chelate effect, optimal spacing between ligand and receptor is not needed.

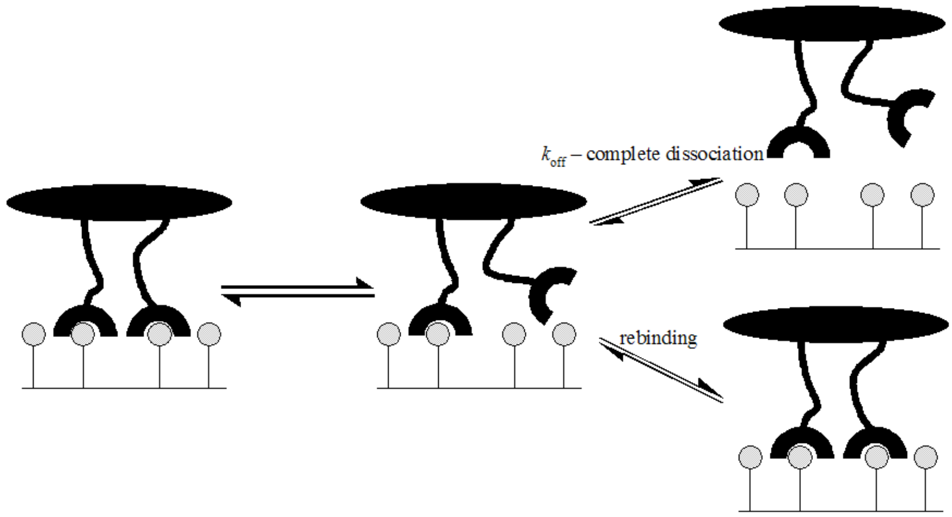


Figure 1.5: The statistical effect

1.4.3 Receptor clustering

Multivalent ligand presentation has been shown to induce cell surface receptors to cluster, thus giving rise to multivalent binding (Figure 1.6). Unlike the chelate effect, in which spatial arrangement of the ligands must be optimal, receptor clustering includes the rearrangement of the receptors into a display set up for the ligands to bind at more than one location. The entropic costs are also different from the chelate effect, as after initial translational costs are taken into account, further interactions do result in further, yet comparatively slight, penalties due to a restriction in receptor mobility. These receptors move in a fluid membrane and thus diffuse in a two- not three-dimensional plane thus limiting the further entropic costs. Receptor clustering events have been shown to be requirements for certain cellular responses (38). One example is the responses from certain toll-like receptors induced by ligand multimerization, which after clustering leads to the recruitment of a variety of adaptor molecules (39).

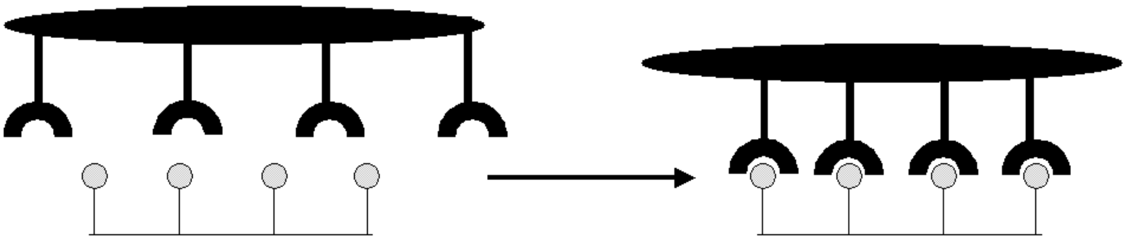


Figure 1.6: Receptor clustering

1.4.4 Subsite Binding

Some lectins have subsites, secondary binding sites, which may be involved in ligand selectivity (40). Ligands may bind to these sites, interacting with a second independent site or with adjacent regions to the primary site (Figure 1.7). Multivalent enhancement can take place when substituents on the ligand interact with extended binded sites via attached residues or parts of the scaffold (41-42). These secondary sites can also be targeted when designing selective inhibitors.

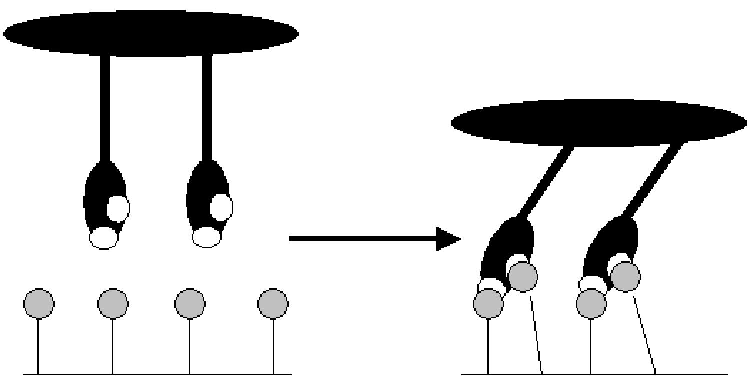


Figure 1.7: Subsite binding

1.4.5 Polyelectric Effects

Many oligosaccharide ligands have anionic moieties including sialic acids and sulfated or phosphorylated sugars. Multivalent scaffolds of charged ligands result in a charge density that can influence binding with either attractive or repulsive electronic interactions. Considering most biological milieus include concentrations of salts, charged ligands can mobilize the counterions in solution, an entropically favorable process (Figure 1.8).

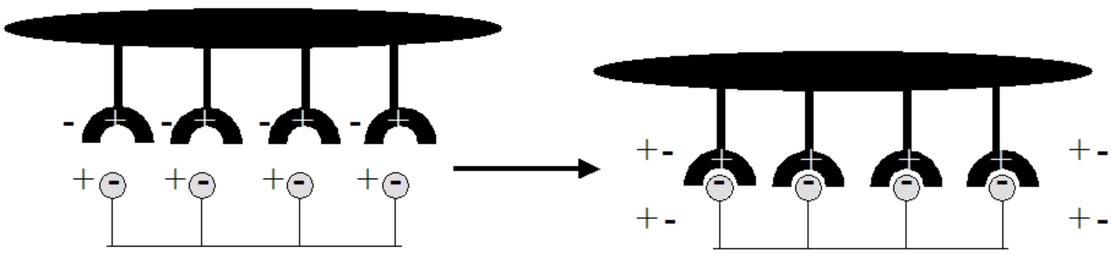


Figure 1.8: The polyelectric effect

1.4.6 Sterics

Although multivalency has been shown to have the effect of greatly increasing the affinity of the interaction involved, not all aspects of multivalency have a positive influence on affinity. Depending on the size of the oligosaccharide and the distance between binding sites on the lectin, multivalent ligands can sterically block the access to neighboring binding sites (Figure 1.9). However, this steric hindrance can be utilized to inhibit pathogens such as viral particles. By the binding of multiple larger species to the protein binding domain, other macromolecules involved in degradation would be obstructed. This idea of steric stabilization is well known in colloidal science but is theorized to work with multivalency as well. This process has been employed in the design of cell-surface inhibitors for glycopolymers on influenza hemagglutinin by Whitesides et. al (43).

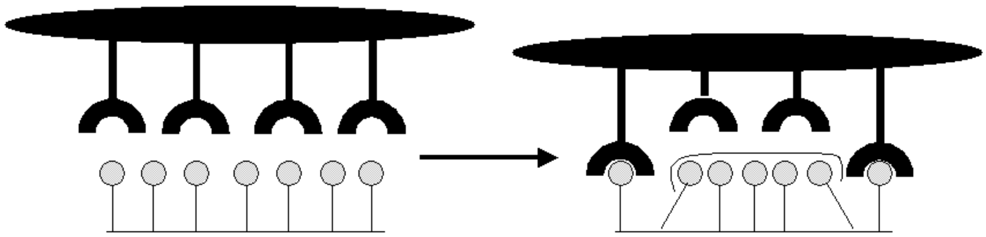


Figure 1.9: Steric hindrance

1.5 Multivalency and Increased Affinity

There are multiple literature examples of multivalent interactions taking place in nature including a large percent of carbohydrate-protein processes (18,44-46). Among these is a study by Lee et al., detailing the extreme increase in affinity possible with the glycoside cluster effect. Lee showed that by increasing the valency of a terminal galactose oligosaccharide from monovalent to divalent to trivalent in a binding experiment with mammalian hepatic asiaglycoprotein, the affinity increased in the ratio of 1 : 1,000 : 1,000,000 (47). Although the mechanism for the phenomenon has never fully been explained, a commonly believed concept is that while once a monovalent bond is lost, the entire interaction is gone, with multiple bonds, a few (or all but one) may be broken, but the interaction as a whole continues until all bonds are broken. This theory holds true for Lee's case as the trivalent compound had a much stronger affinity than the divalent.

Many among the chemistry and biology community have taken the lead from nature by synthesizing multivalent scaffolds to increase both substrate and inhibitor affinity. Synthetic chemists have designed scaffolds containing anywhere from bivalent linkers to multi-generational dendrimers to achieve increased affinities of ligand-receptor bonding. From among these scaffolds, polymers and dendrimers loaded with peripheral sugars, glycopolymers and glycodendrimers respectively, have become two highly utilized classes in analyzing the effects of multivalency. Although vinyl polymerization among other methods leading to polymers and other scaffolds of sugar containing monomers (dendrons etc) is well documented, for the purpose of this thesis coated scaffolds mimicking that mimic oligosaccharides on cell surfaces will be discussed.

Multivalency effects in protein-carbohydrate interactions have been studied expansively using a range of natural and synthetic polymers. Polymers provide a highly valent scaffold without difficult synthesis. Linear polymeric ligands have shown some of the most extreme enhancements of affinity on a valence corrected model. Sialic acid conjugated polymers, reported by Whitesides, Roy and Bovin, have all shown exponential increases in affinity while retaining high aqueous solubility (48-50). Other high valency heterogeneic scaffolds include proteins and gold monolayers (34,51).

Unlike their polymeric counterparts, dendrimers can be prepared as monodisperse homogeneous entities (52-53). These compounds can be synthesized from the periphery inward, convergent synthesis, or from the core out, divergent synthesis, allowing the designer to control physical properties such as size and spatial arrangement. Although dendrimers based on a poly(amidoamine) core, PAMAM dendrimers, have been reported at an eight generation containing over 1000 sugar epitopes, most dendrimers have a much lower valency than that of glycopolymers (54).

1.6 Sialic Acids

Sialic acids encompass a family of naturally occurring 2-keto-3-deoxy-nonic acids known to be involved in an expansive number of biological functions (55-58). Named from the Greek word "sialos" due to the high content in the mucins of saliva, these monosaccharides are often terminal sugars in oligosaccharide chains linked as α-2,3- or α-2,6- to galactosides or 2-acetamido-galactosides and are often located on the surfaces of cells, proteins and other biological molecules (58). There are more than 40 known derivatives of the basic structure (Figure 1.10) (5-Amino-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-ulopyranosonic acid) although the unsubstituted monosaccharide is not found in nature (59). The most abundant natural form, *N*-acetylneuraminic acid

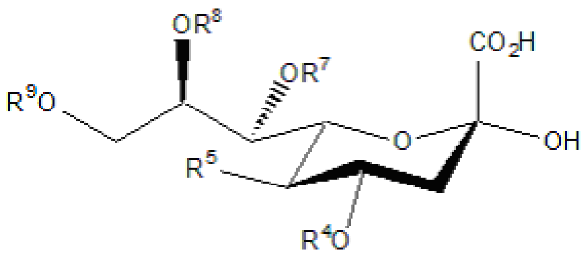
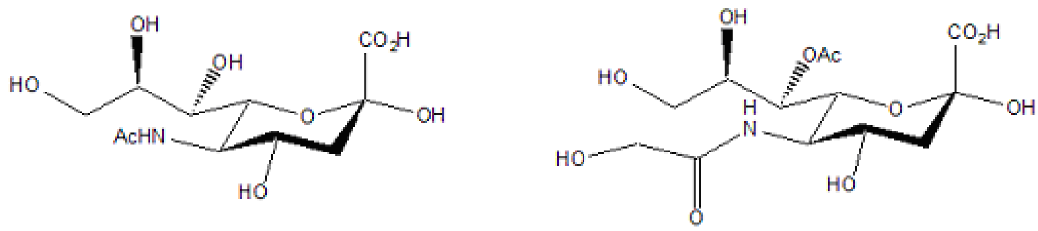


Figure 1.10: Sialic Acid Structure

(Neu5Ac), is characterized by an acetylated *C*-5 amino group (Figure 1.11). Neu5Ac, although non-existent in plants or higher fungi, can be found in the deuterostoma from echinoderms to humans (60-61). Neu5Ac can also be found though rarely in the protostomate lineage as well as in some protozoa, viruses and bacteria (62-64). Substitution of a glycolyl moiety at the same amino group results in *N*-glycolylneuraminic acid (Neu5Gc), another derivative often found in animal species (Figure 1.11). More interestingly, Neu5Gc is not found in humans except in the case of certain cancers (58). Other common modifications include acetylation at *C*-4,7,8 and 9 (59) and lactoylation or phosphorylation at *C*-9 and methylation or sulfation at *C*-8 (65).

The structural diversity inherent in sialic acids is one reason for the multitude of biological functions in which they take part (66). Among these functions include the serving as a ligand for pathogens such as viruses, bacteria and parasites. Sialic acids also serve as recognition sites for certain immune cascades as well as cell adhesion processes (67-68). A major basis for these natural phenomenons is the rather unique structure



αNeu5Ac αNeu5Gc

Figure 1.11: Two Abundant Species of Natural Sialic Acids

of the monosaccharides when compared to most other sugars. The carboxyl group at *C*-1 generates a negative charge on the molecule under physiological conditions making sialic acids suitable for binding and transport of cations and attraction/repulsion events between biological entities (69). When exposed as fields of terminal sugars on surfaces, the negative charge combined with larger size act as a shell protecting the host from various forms of infection (59). Repulsion between multiple molecules containing sialic acids has been known to stabilize the conformation of proteins such as enzymes or lectins (70) as well as preventing coagulation between circulating blood cells (71). The negative charge also ensures the elastic properties of mucins and other protective mucous barriers (72-73). These same chemical characteristics have been utilized to develop analysis methods for sialic acids both natural and derivatized (74).

1.7 Sialidases

Sialidases, synonymous with neuraminidases, are glycohydrolases that catalyze the hydrolysis of the glycosidic linkage between sialic acid and the penultimate sugar of oligosaccharide chains of glycoproteins, glycolipids and other glycoconjugates (59,65,75). All known natural sialidases utilize a double displacement mechanism (Figure 1.12) (76-79) leading to a retention of configuration of the sialoside substrate (76-77). This mechanistic similarity can be accounted for by a conservation of seven amino acids across the entire family that are found to be important in binding and catalysis of the substrate. These residues include an arginine triad which binds to the negatively charged carboxylate group of the sialoside, a glutamate and arginine which form a salt bridge and a tyrosine-glutamic acid diad in conjunction with an aspartic acid, mechanistically involved in the actual catalysis (80-83).

Found in higher order animal species that also contain sialoside substrates, these enzymes fulfill an assortment of roles including internal sialic acid metabolism (59), aging processes of circulating cells (84), apoptosis (85) and immune system responses

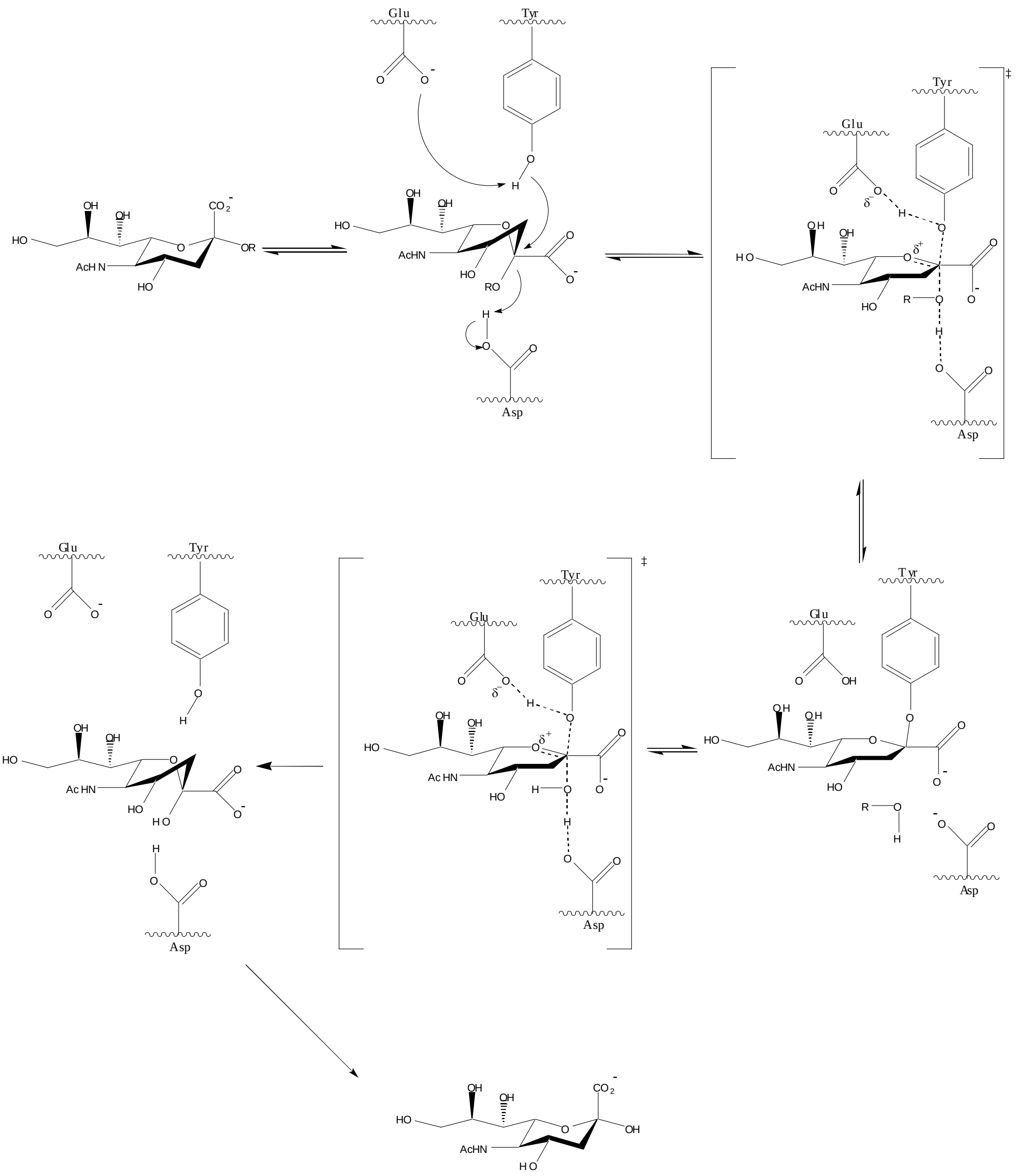


Figure 1.12: Mechanism of sialidase hydrolysis with retention of configuration.

(86). Sialidases have also been found in microorganisms including viruses, bacteria and protozoa, many of which do not generate sialic acid internally (70,87). These often parasitic species, secrete sialidases within a host to scavenge sialic acids for use as a carbon source or as virulence factors that may lead to the onset of disease (59,75). It is the role in disease that has instigated a wave of interest in sialidase's role, mechanism and inhibition.

1.7.1 Viral Sialidases

The earliest implication of sialidase's existence was hypothesized as enzyme activity was shown to release agglutinated influenza virus from chicken erythrocytes by destroying the influenza receptors and were thus originally called receptor-destroying enzymes even before the structure of sialic acid was discovered (88). It is now fully understood that influenza viruses type A and B require sialidases for effective virus mobility as well as replication. After hemagglutinin binds to cell surface terminal sialic acids during initial viral infection, surface sialidases catalyze the cleavage of the sialosides to avoid self –agglutination of the viral buds emanating from the infected cell (Figure 1.13) (89).

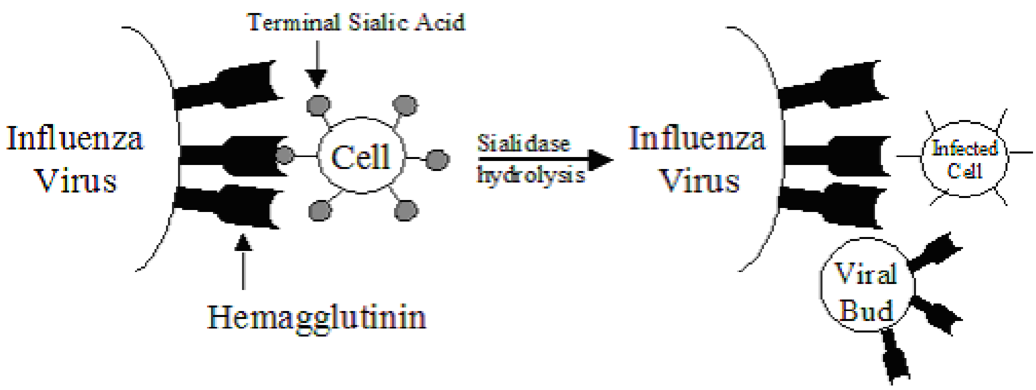


Figure 1.13: Viral sialidase allowing mobility of viral buds

Paramyoxyviruses, including the human type which is a major cause of infant respiratory disease, utilize a similar combination of hemagglutinin and sialidase during infection (75).

1.7.2 Bacterial Sialidases

Sialidases are widely produced by bacterial species. Both pathogenic and nonpathogenic species use these enzymes to fulfill nutritional needs. Sialidases hydrolyze host sialic acids that are then catabolized and transported into the cell for use as both carbon and energy sources (25). In non-pathogenic bacteria strains, many sialidases are involved in symbiotic relationships with the host, often at sialic acid rich mucosal surfaces (25). For example, in the human large intestine, bacterial sialidases preserve an equilibrium to maintain the protective role of the mucous membrane while providing a nutritional source for the entire flora of bacteria necessary for digestion and overall large intestine function (90). A similar symbiosis is seen along human oral mucins as well.

In pathogenic bacteria, the nutritional role is complementary to that of virulence factors in which a range of diseases, both localized and system wide, have been determined (Table 1.1) (25). Although not directly toxic themselves, sialidases released in extreme quantities can cause a variety of detrimental effects via cell damage (91). More often, by altering cell surface environments sialidases improve bacterial pathogenicity by impeding host immunity and/or providing ligands for toxins. One of the best studied occurrences is that of *Vibrio cholerae* removing sialic acid thus freeing the G_{M1} ligand for the cholera toxin to bind (92). Other sialidase activity during bacterial infection includes the destruction of white blood cells, an increase in blood viscosity, an increase in autoantibodies that may lead to autoimmune disease, loss of cell surface negative charge and receptor specificity in the vascular system as well as destruction of mucosal surfaces (25). A commonality of sialidases in pathogenicity is that to take part in disease, the enzyme must be secreted. Nevertheless, although implicated in a number of disease related biological changes, experimental studies determining the necessity of sialidases are somewhat inconclusive.

Table 1.1: Bacterial pathogens and diseases for which sialidases have been implicated as virulence factors.

Microorganism	Disease
Clostridia	Gas gangrene, peritonitis
Streptococcus	Septicaemia, pneumonia, meningitis, glomerulonephritis, periodontal disease
Pneumococcus	Septicaemia, haemolytic-uraemic syndrome
Bacteroides	Septicaemia, peritonitis
Actinomyces	Periodontal disease
Corynebacteria	Septicaemia
Enterococcus	Peritonitis
Escherichia	Peritonitis
Vibrio cholerae	Cholera
Pasteurella	Septicaemia, respiratory disease
Pseudomonas aeruginosa	Cystic fibrosis
Helicobacter pylori	Gastritis

Although sialidase proteins are most often secreted, they have also been found bound to membranes. The latter is the case for some enteric bacteria as well as others including *Antinomyces viscosus* and a number of *Bacteroides* strains (93-95). The function of the cell-bound enzymes is not well understood. However, hypotheses have arisen that they are being stored for later release or their functions are in the host periplasm. Specifically, adherence to epithelial cells is one proposed function for certain cell-bound sialidases (96). Other physical properties including size and pH optimization are species dependent although they generally work at a slightly more acidic range than mammalian lysosomal sialidases which may be a key to inhibiting pathogenic sialidase activity without causing sialidosis, a lysosomal storage disease usually seen as a genetic mutation (25).

1.8 Kinetic Basics

From an atomic level up to that of large biomolecules, interactions leading to a reaction can be studied in turns of kinetics and thermodynamics. While thermodynamics focuses on the energies of reaction, kinetics is concerned with the rate of the reaction. Among the factors influencing the reaction rate include the nature of the reactants in terms of size, electronic effects, pH, etc. Other factors are temperature, the concentrations of reactants, the homo- or heterogenicity of the environment and catalysts, especially enzymes.

1.8.1 Enzyme Kinetics

Enzymes mediate chemical reactions throughout nature. These biocatalysts, while extremely specific on an individual basis, make up a group of several thousand proteins which mediate hydrolysis, oxidation-reduction, dehydration, isomerization and functional group manipulation, among a variety of other reaction types. At the simplest level, enzymes (E) catalyze substrates (S) into one or more products (P). The kinetic study of enzymes was initiated in 1902 when Brown reported the hydrolysis of sucrose by a yeast protein (97) given as:

$$E + S \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P \qquad \{1.2\}$$

As can be seen, enzymatic catalysis requires an initial association step followed by one or more chemical altering steps to give product formation. The initial binding step reaches an equilibrium between the free and bound protein-substrate (ES) complex quantified in terms of the equilibrium dissociation constant, K_d. This constant is determined by the rate of complex association (k_{on}) and the rate of complex dissociation (k_{off}) in terms of enzyme, substrate and complex concentrations:

$$K_d = \frac{k_{off}}{k_{on}} = \frac{[E]\ [S]}{[ES]} \qquad \{1.3\}$$

The overall rate for the collective steps after complex equilibrium is termed k_{cat}. When real-time kinetic determination cannot be accomplished, kinetic rates can be determined via the *steady state* approach of Briggs and Haldane, assuming rate equilibrium is reached, that is [ES] remains approximately constant from milliseconds after mixing until the substrate is nearly expended (98):

$$\frac{d[\mathrm{ES}]}{dt} = 0 \qquad \{1.4\}$$

Where equation 1.4 is true, the reaction velocity (v) can be determined by:

$$v = \frac{V_{\mathrm{max}}\,[\mathrm{S}]}{K_{\mathrm{m}} + [\mathrm{S}]} \qquad \{1.5\}$$

with V_{max} equal to the maximum reaction velocity defined as:

$$V_{\mathrm{max}} = k_{\mathrm{cat}}\,[\mathrm{E}] \qquad \{1.6\}$$

and K_{m}, identified as the *Michaelis constant*, and is equal to the substrate concentration giving half V_{max} under saturating conditions:

$$K_{\mathrm{m}} = \frac{k_{\mathrm{off}} + k_{\mathrm{cat}}}{k_{\mathrm{on}}} \qquad \{1.7\}$$

K_{m} and K_{d}, although not equal, become the same constant when applying the models of Henri (99) and Michaelis and Menten (100). The assumption of this model requires a rapid equilibrium between E, S and ES followed by the rate-determining step of the conversion of ES to E and P. Thus $k_{\mathrm{cat}} << k_{\mathrm{off}}$ so $k_{\mathrm{off}} + k_{\mathrm{cat}}$ approaches k_{off} and thus equation 1.7 becomes equal to equation 1.3 making $K_{\mathrm{m}} = K_{\mathrm{d}}$.

Perhaps the best way to measure an enzyme's catalyzing ability and to compare multiple enzymes is through the catalytic efficiency, k_{cat}/K_m. From equation 1.6 and the Michaelis constant, the catalytic efficiency can be calculated as;

$$\text{catalytic efficiency} = \frac{V_{max}}{K_m\,[E]} \qquad \{1.8\}$$

In order to determine the catalytic efficiency, the reaction must effectively reach a state of a bimolecular reaction between free enzyme and free substrate or $[S] <<< K_m$..

1.9 Non-enzymatic binding

As discussed earlier, enzyme-substrate interactions are only one type of protein-analyte associations. Other types utilizing a receptor-ligand formation, initialize biological communication, immunological cascades as well as simple adhesion. Compared to enzymes that catalyze a reaction giving a product, these interactions have an on/off component without substrate transformation:

$$R + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} RL \qquad \{1.9\}$$

where enzyme (E) and substrate (S) have been replaced by receptor (R) and ligand (L).

1.10 Pattern Recognition Recptors

Among the receptor-ligand type biomolecular associations, there are numerous protein receptors involved in the active defense mechanisms against invading pathogens. These proteins, key components in innate immunity, the first line of defense against microorganisms in vertebrates and the only defense for invertebrates and plants, recognize conserved motifs not found in higher eukaryotes (101-103). When present on cell surfaces, these proteins are referred to as pattern recognition receptors (PRRs), which directly detect pathogen-associated molecular patterns (PAMPs) to initiate immune responses (103). PAMPs, which are not restricted to pathogenic microbes, include lipopolysaccharide (LPS) found in the cell wall of Gram-negative bacteria, peptidoglycan (PGN) common to almost all bacteria, bacterial flagella, yeast cell wall components and double-stranded RNA from viruses (104-105). The PRRs that recognize these PAMPs and induce host responses include toll-like receptors (TLRs), nucleotide-binding oligomerisation domain (Nod) molecules, peptidoglycan recognition proteins (PGRPs) as well as possibly CD14 and others.

1.10.1 Toll-like Receptors

Originally discovered in *Drosophila melanogaster* as a protein involved in embryogenic development (106), Toll was later associated with the immune system when the deficiency of this protein in flies led to much increased susceptibility to fungal and Gram-positive bacterial infections (107-108). In 1998, Rock et al. discovered the first Toll-related protein in mammals (109). Soon after, it was verified that TLR4 was the signaling receptor in a complex cellular response to LPS recognition (110). It is now known that TLRs recognize other PAMPs, for example lipotichoic acid from Gram-positive bacteria, double-stranded viral RNA and flagellin (105,109).

Toll-like receptors are transmembrane proteins that are members of the interleukin-1 receptor (IL-1R) superfamily characterized by a leucine-rich repeat unit in the extracellular region and a Toll/IL-1R (TIR) cytoplasmic domain (111-112). Once a PAMP binds to the extracellular domain, the intracellular TIR recruits adapter molecules such as MyD88 to further mobilize a cascade of cell signaling events leading to a pro-inflammatory response (113-114). As of now, 11 TLRs have been identified via structural homology, although the functions of all have not yet been established (115).

1.10.2 Nod proteins

Nod1 and Nod2 comprise another family of PRRs. Discovered to be involved in intracellular pathogen detection, these cytosolic proteins are comprised of a *C*-terminal series of leucine-rich repeats and a central nucleotide binding domain. Nod1 has a caspase-activating and recruitment domain (CARD) at the *N*-terminus whereas Nod2 has two of these domains (104). CARD domains play an integral role in protein-protein interactions.

Both Nod proteins recognize the PAMP peptidoglycan although the substructure required is protein dependent. Nod1 detects a naturally occurring degradation product of PGN requiring an exposed *meso*-diaminopimelic acid residue found in Gram-negative bacteria, but very few strains of Gram-positive bacteria (116). Nod2 on the other hand recognizes MDP (see chapter 3), a motif found in both Gram-positive and Gram-negative bacteria (117-118).

Since Nod proteins are intracellular, PGN must find a way into the cell before recognition and subsequent immune activation can occur. Although MDP has adjuvant activity, it is likely that a co-factor is involved such as lipid modification to enhance activity (104). PGN may also gain cell entry through bacterial infection as shown by Girardin's study of *Shigella flexneri* (119). Although yet to be proven, phagocytosis may be another vehicle for the internalization of PGN for Nod activation (104).

1.10.3 Peptidoglycan Recognition Proteins

First discovered by Ashida and co-workers when the protein from *Bombyx mori* bound PGN activating the prophenoloxidase immune cascade (120), PGRPs are now know to be conserved from insects up to higher eukaryotes including mammals (121). PGRPs include extracellular, intracellular and transmembrane proteins divided into short (PGRP-S) and long (PGRP-L) depending on size. As with other PRRs, PGRPs function are related to innate immunity.

The 17 known PGRP proteins of *Drosophila* as well as a number of their immune functions are well documented. As with the silk worm, PGRP-LE from *Drosophila* activates the prophenoloxidase cascade (122), an immune pathway found only in insects. Reception of PGN from the Lysine-type PGN usually found in Gram-positive bacteria activates the Toll pathway (123). The *Drosophila* Toll-independent *Imd* pathway is activated by Gram-negative bacteria as well as some Gram-positive bacilli through PGRP-LC (124-126). This pathway is similar to the TNF-α receptor-induced pathway found in mammals, thus increasing the interest in the four mammalian PGRPs (127-128).

However, at this time, less is know about the mammalian PGRPs, especially in humans (see chapter 3).

1.10.4 CD14

CD14 is a cell-surface protein under intense scrutiny due to a controversy on whether it is a PRR. The ability to recognize specific shared features of microbial cells while discriminating between ligands to control immune response, a requirement of PRRs, is under dispute due to the inability of CD14 to differentiate agonistic and antagonistic versions of LPS (129-130). For this thesis, a cautious view of CD14 as a PRR is taken.

CD14 is a glycosylphosphatidylinositol-linked (GPI) protein that acts as a receptor for LPS (131-132), PGN (133) and possibly bacterial glycopolymers (134-135) and lipoteichoic acid (136). LPS or PGN binding activates a pathway that stimulates macrophage release as well as the secretion of mediators (137). CD14 is also found in soluble form in normal serum and milk (137). As a soluble protein, CD14 complexes with LPS to activate CD14-negative cells (138). The soluble form also enhances the response of CD14-positive cells to LPS and PGN (139).

1.11 Surface plasmon resonance: Basics

Biophysical binding assays that produce high-quality label-free data on the interactions between protein targets and potential drug candidates have become indispensable tools in the drug discovery process. Optical biosensors based on surface plasmon resonance (SPR) technology have the ability to monitor reversible binding associations of label-free biological molecules in real time. Binding assays based on SPR have several advantages compared to other analysis methods, including low target consumption, simultaneous referencing and the ability to resolve associations into *on* and *off* rate constants. These advantages have made SPR a key tool in determining the affinity, kinetics and thermodynamics of a number of biological systems.

Surface plasmons are oscillating collections of free electrons that occur in abundance at the surface of metals such as aluminum, silver and gold (140). The properties of the interaction of polarized light at the surface of thin metal films at which these plasmons are created are the basis for SPR technology (141). By utilizing a prism at the metal interface, a transitory wave, present in total reflection, causes an optical excitation of the surface plasma wave that can be seen as a decrease in reflection for a unique angle of incidence (142). The amplitude of this change in the angle of incidence is proportional to a change in the refractive index of the layer contacting the metal film opposite to the prism. Binding of the substrate of interest at this layer increases the mass and thus the refractive index proportionally. Therefore, the binding can be analyzed quantitatively resulting in the desired kinetic data (Figure 1.14).

In order to achieve the obtain the desired biophysical phenomena, it is compulsory to immobilize one of the two interacting biological molecules to a surface coated with the necessary metal film. Biacore AB (Uppsala, Sweden), the leader in commercial optical bionsensors based on SPR technology, has developed an array of sensor chips on which the immobilization and subsequent experimentation can take place.

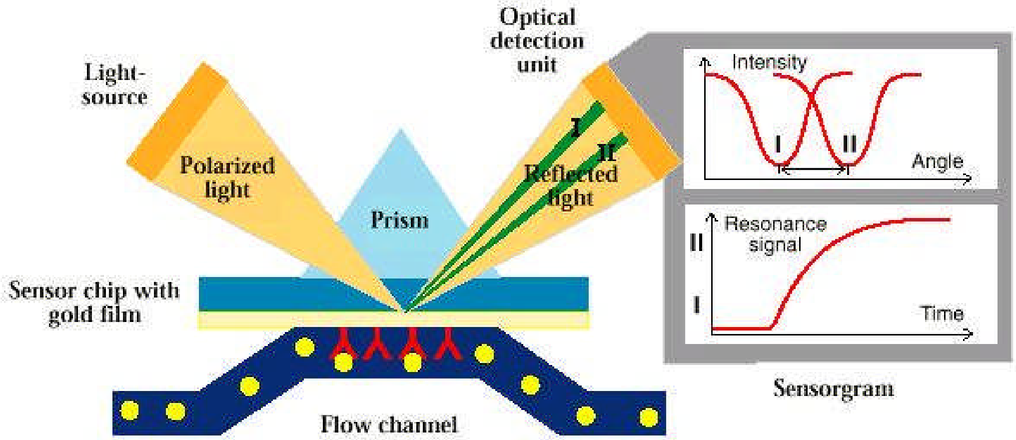


Figure 1.14: The basis for surface plasmon resonance technology (143).

The most basic sensor chip consists of a glass slide coated with a thin layer of gold, to which a carboxymethylated dextran matrix is covalently attached (Figure 1.15). Immobilization of biomolecules that have amino functional groups can be achieved by a succinimide activation of the carboxyl groups on the CM5 sensor chip followed by amide bond formation (Figure 1.16). Although this technique is the most widely applicable due to the majority of immobilizations utilizing proteins, which often have residues with side chain amino functionalities such as lysines, other sensor chips and consequent

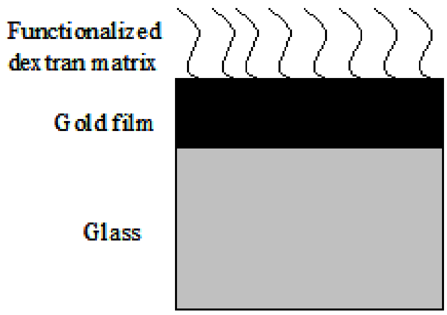


Figure 1.15: Biacore sensor chip design.

immobilization approaches including thiol coupling through cysteine residues and capture techniques utilizing the strong non-covalent biotin-streptavidin interaction, are also quite common. In order to optimize response depending on the type of application desired, different ligand densities can be achieved via concentration and time dependent immobilization procedures. For example, a low density is more advantageous for kinetic experiments to minimize steric and rebinding issues. On the other hand, higher loaded surfaces are beneficial for concentration measurements.

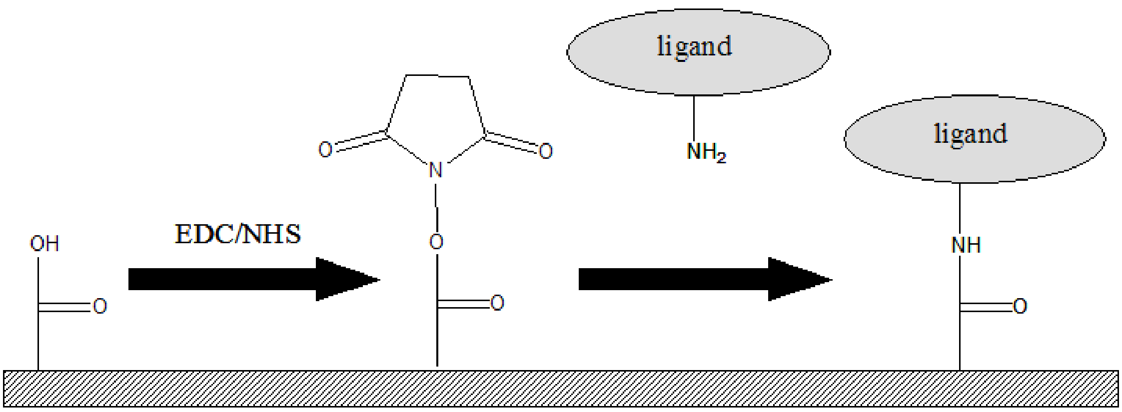


Figure 1.16: Amine coupling immobilization of the ligand to the sensor chip.

1.12 Analytical Determination Using SPR

Depending on this level of immobilization, a theoretical maximum response termed R_{max} can be calculated describing the binding capacity of the surface at saturation using the formula (143):

$$R_{max} = (\text{analyte MW})/(\text{ligand MW}) \times \text{ligand density} \times \text{stoichiometric ratio} \qquad \{1.10\}$$

The theoretical R_{max} is often slightly higher than the experimentally derived R_{max} due to steric interactions and the partial inactivity of the ligand.

Once the ligand of interest is immobilized onto the chip, the second biological molecule can be flowed over the chip surface (Figure 1.17). As binding takes

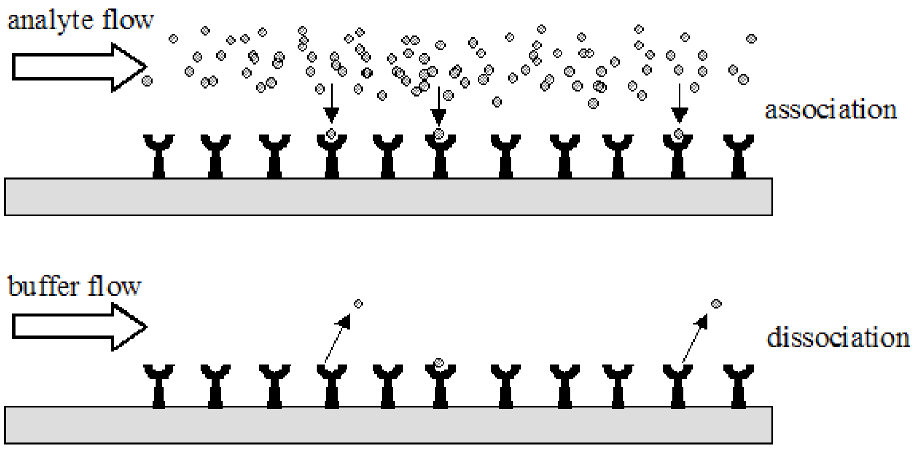


Figure 1.17: Schematic representation of the association and dissociation phases of an optical biosensor experiment.

place, the mass and thus the refractive index at the surface increases. At a given time, the analyte flow is terminated and the bound analyte freely dissociate. Remaining bound molecules can be removed using regeneration buffers, thus rendering the surface refreshed for the next run. The change in refractive index, recorded in response units (RU), is charted versus time on a real time scale as a sensorgram (Figure 1.18) including associative and dissociative phases that can be analyzed to determine the kinetics of the biomolecular interaction.

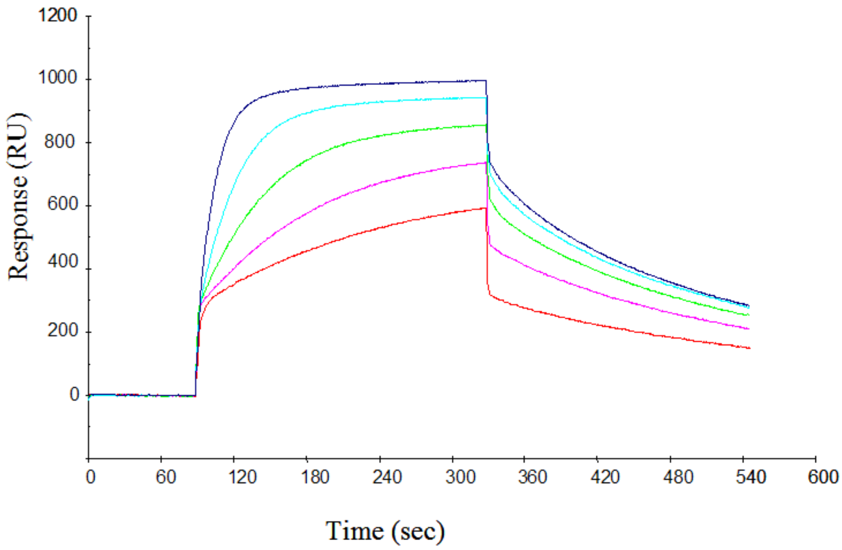


Figure 1.18: Typical sensorgram for a range of concentrations of analyte.

For a classical bimolecular interaction, the Langmuir equation (144) can be written as (145):

$$K(n-r) = rg^{-1} \qquad \{1.11\}$$

where K is the affinity constant in l/mol, n is the total sites on the chip, r is the analyte occupied sites and g is the concentration of the analyte in mol l^{-1}. From the response (R) calculated by the instrument, the equilibrium value of K_a can be calculated from

$$R_{max} = R_0 + (R_{sat} - R_0)\, K_a\, g_0\, (1 + K_a\, g_0)^{-1} \qquad \{1.12\}$$

where for a given concentration of ligand g_0, R_{max} is the maximum value of R at $t = \infty$ interpreted from the curve

$$dR/dt \text{ versus } R \qquad \{1.13\}$$

R_0 is equal to R during normal ligand-free buffer flow and R_{sat} is the value of R at saturation of the binding sites. From the experimental curves, the set of kinetic parameters can be determined (146).

During the association phase, R increases with respect to

$$R = R_0 + (R_{max} – R_0)\, [1 – e^{(k_{bind} t)}] \qquad \{1.14\}$$

where k_{bind} is the experimental kinetic binding constant

$$k_{bind} = k_{ass} g_0 + k_{diss} \qquad \{1.15\}$$

and k_{ass} is the kinetic association constant in l mol^{-1} s^{-1}, g_0 is the concentration of free ligand in mol l^{-1} and k_{diss} is the kinetic dissociation constant in s^{-1}. During the dissociation phase in which ligand-free buffer flows across the cell, R decreases by

$$R = R_\infty + (R_{max} - R_0)\,[1 - e^{(k_{diss} t)}] \qquad \{1.16\}$$

where R_0 is R after complete dissociation of the analyte.

The data can be analyzed with linear transformations or according to the integrated rate equations describing the association and dissociation of the analyte to the ligand. Direct association and dissociation rate constants were shown for singular binding experiments using non-linear least squares methods by O'Shannessy et al. (147). Although these rate equations allow the kinetic analysis of most experimental data, due to the inability to completely purify certain biological milieus as well as instrumental limitations, certain more common pitfalls should be examined.

Although the analyte flow is directed across the cell surface, laminar flow leaves a small layer of stationary buffer between the analyte and ligand (143). Thus the final distance before interaction takes place is diffusion controlled. There are certain instances when this diffusion is limiting and steps should be taken to avoid them. Molecular weight and diffusion constants are proportional thus methodology for higher weight analytes should be carefully planned. Faster analyte flow lessens the stationary layer between the analyte and ligand, thus diminishing the diffusion time. Consequently, low flow rates should be avoided, especially with large molecular weight analytes, unless analyte quantity is severely limited. Smaller concentrations of analyte reduce mass transfer effect as well so systems should be kept well under saturation when possible. Confirmation of mass transfer effects can be qualitatively measured by a plot of dR/dt versus R. Negative curvature at concentrations higher than the reciprocal of the affinity constant usually signifies diffusion limitations and in such cases, optimization of methodology must be revised (148).

Often, one or both of the biomolecular moieties analyzed in a SPR experiment are obtained from the purification of a biological sample. Non-complete purifications may cause deviation from the pseudo-first order kinetic behavior assumed by the rate equations due to heterogeneity in either the ligand or the analyte. Some proteins have multiple yet unequal binding domains which may bind at different rates and/or concentrations of analyte. These biphasic binding patterns can be seen qualitatively by a deviation in the slope of the real time sensorgram. An in-depth view of the possible causes of these deviations is reviewed by O'Shannessy and Winzor (149). Although these complications can rarely be fully resolved, sensorgrams reaching near equilibrium can be utilized to obtain steady state data (Figure 1.19) (150).

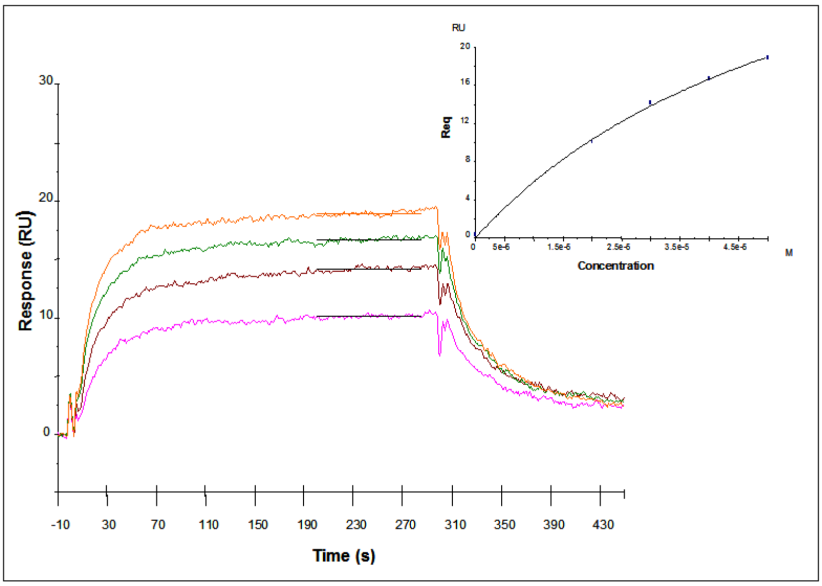


Figure 1.19: Sensorgram utilizing equilibrium based analysis to derive steady state constants (151).

1.13 Research Projects

Over the past few decades, the importance in carbohydrate recognition and biological response has become apparent. As a result, the comprehension of the kinetics of initial binding, cell signaling, inhibition and all other biomolecular associations involving glycoconjugates has become vital in understanding the phenomena in which they play a role. Via synthetic glycoconjugates, the study of these small molecule-protein interactions are becoming resolved.

1.13.1 Mechanism and Inhibition of Modular Bacterial Sialidases

There are no high affinity inhibitors for bacterial sialidases to date. Additionally, nonselective inhibitors of sialidases may impede human sialidases causing toxic side effects such as those caused by sialidosis. In order to design a highly potent while selective inhibitor of bacterial sialidases, the complete understanding of the nature of these proteins is necessary.

First, it was necessary to synthesize a substrate in a quantity enough to satisfy all of the kinetic and inhibitory experiments required during the project. Thus the sialyl lactosamine trisaccharide from the non-fucosylated sialyl Lewis^{X} was chosen (Figure 1.20). To replicate a better mimic of the substrate, conjugation to a polymer was utilized to copy the cell surface aggregation of these ligands in nature (Figure 1.21).

The kinetics of the synthetic substrates (monovalent and polyvalent) were studied utilizing a range of modular bacterial sialidases as well as a control bacterial sialidase not having lectin domains. To verify the new mechanism discovered, a range of inhibitors were required. Due to obstacles making the previous assay unviable, a new assay was

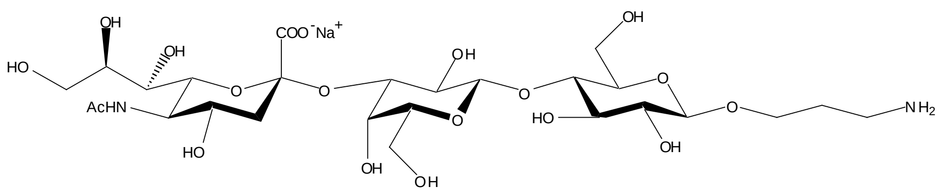


Figure 1.20: Monovalent Substrate

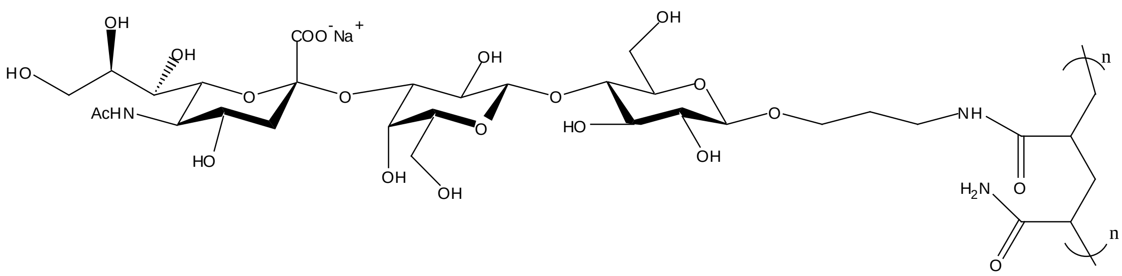


Figure 1.21: Polyvalent Substrate

developed and all previous kinetic studies were re-examined. Utilizing the new mechanism for modular bacterial sialidases, new selective inhibitors were designed, synthesized and kinetically studied using the new assay (Figure 1.22).

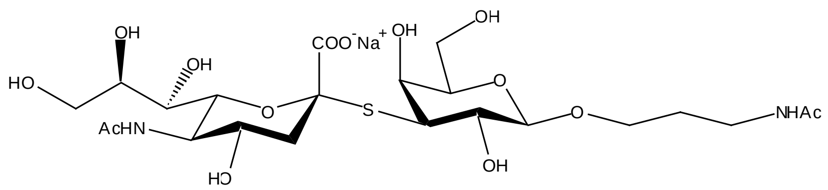


Figure 1.22: Thio Inhibitor

1.13.2 Ligand Specificity for human Peptidoglcyan Recognition Proteins

PGRPs are vital proteins in recognition and initial innate immune responses in insect biology. Although some testing has occurred with mouse and bovine PGRPs, little is known about the four human PGRPs. Taking advantage of surface plasmon resonance technology and a family of peptidoglycan part structures synthesized in the Boons laboratory, the binding affinities and selectivities of hPGRP-S and hPGRP-Iα were studied.

Although SPR has become a powerful tool in elucidating the kinetic interactions between biomolecules over the past decade, little work has been accomplished when using small molecule analytes. Due to the increased risk that manipulations to a small molecule could have overall effects that may change the nature of the binding interaction, immobilizing small molecules onto the chip surface is not always advantageous. Thus, pioneering work in small molecule analyte assays for SPR are required to test the PGN part structures.

Chapter 2

Mechanistic Studies and Selective Inhibition of Modular Bacterial Sialidases

2.1 Introduction

Bacterial sialidases act as virulence factors in a variety of important diseases including cholera, meningitis, septicaemia and cystic fibrosis (152). Some specific roles of sialic acid hydrolysis in pathogenesis include the secreted sialidase from *Vibrio cholera* desialylating higher order oligosaccharides to reveal the GM1 receptor for cholera toxin (153) and the provision of nutritional sources for the pathogen by providing both a carbon and energy source (154). Inhibition of these pathogenic sialidases is a key to diminishing and possibly deleting the virulence of these organisms. However, sialidase function in mammalia is necessary to regulate surface sialic acid content of cells required in such functions as immune response and signal transduction (155-156) and the diminished efficiency of these enzymes has been implicated in certain lysosomal storage diseases such as sialidosis (157). Yet to date, the ability to design high affinity inhibitors specific for pathogenic sialidases has been primarily unsuccessful. Among the pathogenic bacteria of interest include those employing modular sialidases such as *Vibrio cholera* and *Clostridium perfringens* (158).

Over the past decade, the known abundance and variety of complex modular proteins have greatly increased making them a viable subset for thorough studies (159). Initially, carbohydrate binding domains, as well as other lectin-type modules bound to primary proteins, were hypothesized to be recognition markers for cell surface oligosaccharide substrates (160). Further experimentation has shown cellulose binding domains in certain cellulose hydrolyzing enzymes can actually interfere with cellulose self-adhesion within plant cell walls and thus aids in the enzymatic breakdown process (160). More recently, we uncovered certain modular bacterial sialidases that increase catalytic efficiency by utilization of covalently bound carbohydrate binding domains (161). Through multiple simultaneous interactions between these enzymes and the oligosaccharide ligand clusters on cell surfaces, the affinity of these biomolecules is greatly increased leading to a much greater normalized catalytic efficiency of hydrolysis as compared to a single ligand-single domain model. Enzymatic catalysis thus becomes a new example of multivalency in biological function.

Multivalency in nature has widespread applications as it is an efficient way to increase poor affinity interactions without the restructuring of either protein or ligand (162-167). Carbohydrates often show a relatively weak affinity for the receptors they interact with and are thus prime examples of the employment of multivalency (162). Many well distinguished carbohydrate-protein interactions utilize this strategy to increase affinity (162-164), change selectivity (165) or intiate cell signaling (166-167). As such, carbohydrates are often found clustered on cell surfaces to greater facilitate multivalency including oligosaccharides terminating in sialic acid (168).

Neuraminic acid (sialic acid) is an abundant, often terminal, sugar in many glycoproteins and glycolipids found throughout biological systems (169). Sialidases catalyze the hydrolysis of these sialic acids (152), often as an initiation point for harmful microorganisms to invade (153). Due to their involvement in a wide range of diseases (170), the ability to test the kinetic efficiency of these enzymes, and furthermore, the inhibitory activity of molecules designed to slow or stop catalysis is crucial. Typically, sialidase catalysis has been assayed by thiobarbituric acid (TBA) method (171), radio labeling,(172) and fluorescent labeling (173). Each of these methods has their advantages as well as drawbacks. The TBA assay, although relatively simple, lacks in sensitivity. More concerning is the inability of TBA to test numerous inhibitors due to secondary interactions with the post enzymatic chemical reactions necessary to test for free sialic acid. Radioisotopic assays are quite sensitive but require the handling and disposal of radioactive materials throughout the synthesis and kinetic applications of these substrates. Fluorescence is also a very sensitive method, but pre-catalysis labeling, such as methods using 4-methylumberipheryl derivatives (174), require time consuming preparations and evaluate unnatural compounds. Fluorescent assays which label after hydrolysis using 1,2-diamino-4,5-methylenedioxybenzene (175) avoid these encumbrances, but as with TBA, can react with inhibitors to give false analysis.

2.2 Mode-of-Action Studies

While determining the effects of multivalent substrates on modular bacterial sialidases, we uncovered a new example of multivalency. This novel mechanism, rationalized by a model in which the catalytic and lectin domains of a modular enzyme interacts simultaneously with a polyvalent substrate, gives rise to a prominent increase in catalytic efficiency for polyvalent substrates versus their monovalent counterparts on a per mole basis. This mode-of-action was utilized in the design of the first polyvalent inhibitor of *Vibrio cholerae* targeting the lectin domain and thus not based on a sialic acid related scaffold. This inhibitor design also reveals an uncomplicated synthetic pathway to selectively inhibit modular enzymes containing a catalytic domain and one or more binding domains.

2.3 Results of Mode-of-Action Studies

In order to study the effects of substrate clustering on the hydrolyzing ability by the modular sialidases from *Clostridium perfringens* and *Vibrio cholerae* (170,176) as well as the sialidase from *Salmonella typhimurium* which does not contain lectin domains (177), monovalent trisaccharide 2.1 and glycopolymer 2.2 (Figure 2.1) were synthesized. Enzymatic hydrolysis was achieved while varying the substrate concentration corrected for valency. The rate of reaction was determined by quantifying freed sialic acid using an HPAEC-based method while the Michaelis constants (K_m) and maximum velocities (V_{max}) were determined from the nonlinear fitting method of the Michaelis-Menten equation. To better compare velocity measurements, the V_{max} data were calculated relative to one another with the monovalent compound arbitrarily set to 1.

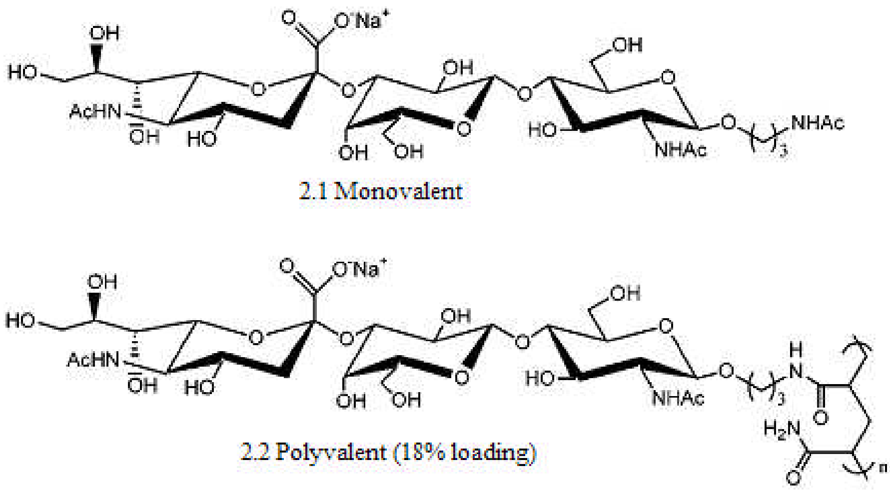


Figure 2.1: Mono- and polyvalent 3'-sialyl N-acetyllactosamine

The apparent kinetic parameters obtained from the two substrates demonstrated that polyvalent 2.2 was hydrolyzed much more efficiently than monovalent 2.1. Both modular enzymes displayed K_m values approximately 100-fold smaller for the polyvalent trisaccharide 2.2, while V_{max} showed a minor decrease to give an overall significant increase in catalytic efficiency (V_{max}/K_m) (Table 2.1). However, the non-modular *Salmonella typhimurium* displayed no enhancement of hydrolysis for either substrate, implying that the binding domains may cause the increase in efficiency.

Table 2.1: Apparent Kinetic Parameters of Hydrolysis of Monovalent 2.1 and Polyvalent 2.2 by Sialidases (161)

Sialidase	Cmpd	K_m (mmol)	rel V_{max}	V_{max}/K_m
Clostridium perfringens	2.1	2.2 ± 0.3	1 ± 0.10	0.45 ± 0.07
	2.2	0.04 ± 0.02	0.85 ± 0.13	21 ± 10
Vibrio cholerae	2.1	5.7 ± 0.4	1 ± 0.09	0.18 ± 0.02
	2.2	0.04 ± 0.01	0.23 ± 0.02	5.8 ± 1.4
Salmonella typhimurium	2.1	2.1 ± 0.5	1 ± 0.11	0.48 ± 0.13
	2.2	2.9 ± 0.3	1.4 ± 0.2	0.48 ± 0.06

Initially, we devised a hypothesis in which a high local substrate concentration surrounding the polymer scaffold caused the increase in activity. A low loaded polymer (2%) had a K_m of 1.5 mM with *Clostridium perfringens*, strengthening the idea of a local concentration effect. A control experiment using a co-polymer loaded with 2%

trisaccharide substrate and 16% melibiose as an inactive dummy sugar was run to negate the possibility that polymer shape, and thus activity, was dependent on loading. The surprising result with the *Cp* enzyme gave a K_m of 0.06 mM, much lower than would be expected by our theory. It was this control experiment with the galactose terminal melibiose that led to further investigations of modular enzymes and the novel mechanism of multivalency increasing catalytic efficiency.

To further the rationale that the increased efficiency, shown to be caused by an increase in affinity for the polyvalent substrate, was due to a multivalency effect of the lectin domain(s) binding to the substrate simultaneously with the catalytic domain, we formulated an experiment in which inhibition of the lectin domain would delete the increased efficiency, leaving the polyvalent and monovalent substrates with similar kinetic parameters. D-Galactose was chosen to examine this effect due to terminal sialic acids often, including the case of trisaccharides 2.1 and 2.2, being linked to galactosides. Furthermore, an X-ray crystal structure of another modular bacterial sialidase, *Micromonospora viridifaciens*, shows a complexation of the binding domain with D-Galactose (158). The proposed experiment agreed with the suggested mechanism as an increase in free D-Galactose as an inhibitor increased the K_m, and thus lowered the catalytic efficiency, of *Vibrio cholerae* towards polyvalent 2.2 (Table 2.2) while not having an effect with monovalent 2.1 or with the sialidase from *Salmonella typhimurium* indicating that D-Galactose does not interfere with the catalytic domain.

Table 2.2: Apparent Kinetic Parameters of *Vibrio cholerae* with D-Galactose Inhibition (161)

Substrate	[I] (mmol)	K_m (mmol)	rel V_{max}	V_{max}/K_m
2.1	0	5.7 ± 0.4	1 ± 0.09	0.18 ± 0.02
2.1	5	5.3 ± 0.4	0.80 ± 0.05	0.15 ± 0.01
2.2	0	0.04 ± 0.01	0.23 ± 0.02	5.8 ± 1.4
2.2	5	0.09 ± 0.02	0.26 ± 0.02	2.9 ± 0.6
2.2	10	0.20 ± 0.06	0.33 ± 0.02	1.6 ± 0.5
2.2	15	0.5 ± 0.1	0.43 ± 0.02	0.86 ± 0.17

Under conditions in which the environment is saturated with product, sialidases can catalyze trans-sialylation, the transfer of sialic acid to a galactoside (178). The inhibition assay was repeated with D-Lactose utilizing sialyllactose as an HPLC standard to rule out trans-sialylation as a contrary solution to the observed effects. Only 5% of the trans-sialylation product was observed relative to hydrolyzed sialic acid and was corrected. The new results showed a minor increase in V_{max}, but no change in K_m.

Further analysis of the inhibition of *Vibrio cholerae* by D-Galactose (Dixon plot analyzed) reveals competitive inhibition at low to moderate substrate concentrations with a K_i = 5 mmol (Figure 2.2). At high substrate concentrations, the increased affinity toward the substrate will actually slow hydrolysis due to slightly declining V_{max} values. At these substrate levels, D-Galactose leads to slight activation. However, these concentrations are well above what would ever be seen in a biological environment.

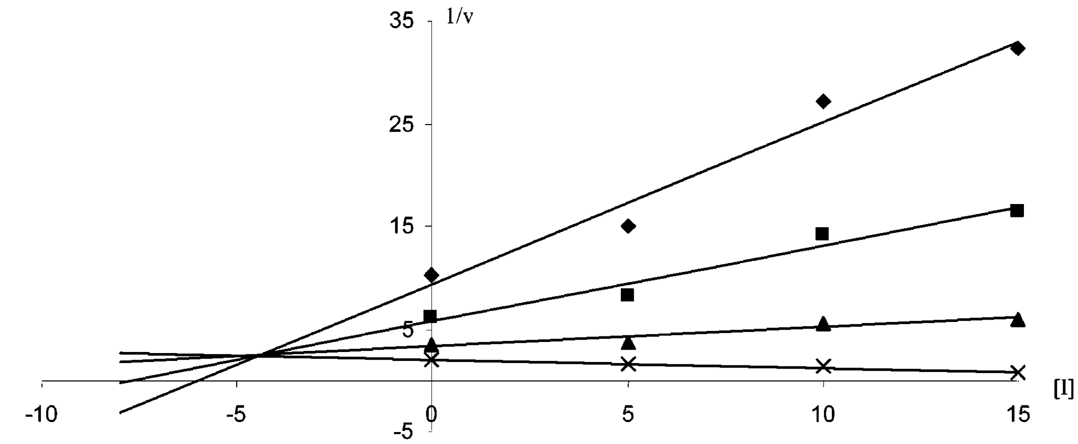


Figure 2.2: Dixon plot of hydrolysis by *Vibrio cholerae* of polyvalent 2.2 with different concentrations of D-Galactose; ◆,0.01; ■,0.02; ▲,0.06; and ×, 4.0 mmol (valency-corrected) (161).

Vibrio cholerae contains two carbohydrate binding domains flanking the catalytic fold and thus it was queried whether a polyvalent inhibitor would be able to achieve increased avidity by interacting with both lectin domains simultaneously. A galactose terminal sugar, aminopropyl melibiose [Gal(1-6)Gal-$(CH_2)_3NH_2$] (Figure 2.3) (synthesized elsewhere in the Boons' lab) was conjugated to a poly[n-acrylamide] backbone to test whether a polyvalent inhibitor would show increased inhibition. As with D-Galactose, increasing concentrations of the polyvalent inhibitor resulted in an increase in K_m and a decrease in catalytic efficiency (Table 2.3). Dixon plot analysis gave a valency corrected K_i = 50 µmol, making the polyvalent compound a 100-fold better inhibitor than the monovalent D-Galactose (Figure 2.4). As with D-Galactose, the

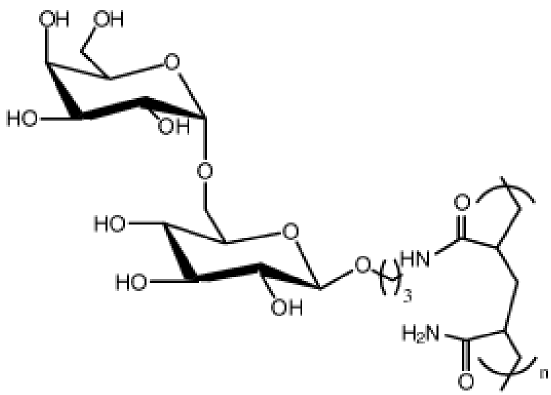


Figure 2.3: Aminopropyl melibiose conjugated polymer (18% loading)

polymeric inhibitor did not affect *Salmonella typhimurium.* This polyvalent inhibitor provides a significant step in the development of inhibitors for bacterial sialidases for which the development of potent inhibitors has been challenging (179), while also

Table 2.3: Apparent Kinetic Parameters of *Vibrio cholerae* with Polyvalent Melibiose Inhibition (161)

Substrate	[I] (mmol)	K_m (mmol)	rel V_{max}	V_{max}/K_m
2.2	0.10	0.25 ± 0.04	0.48 ± 0.07	1.9 ± 0.3
2.2	0.25	0.54 ± 0.1	0.81 ± 0.06	1.5 ± 0.3
2.2	0.50	1.6 ± 0.3	1.1 ± 0.2	0.69 ± 0.16

providing the ability to selectively inhibit modular sialidases versus those containing catalytic domains only (all human known).

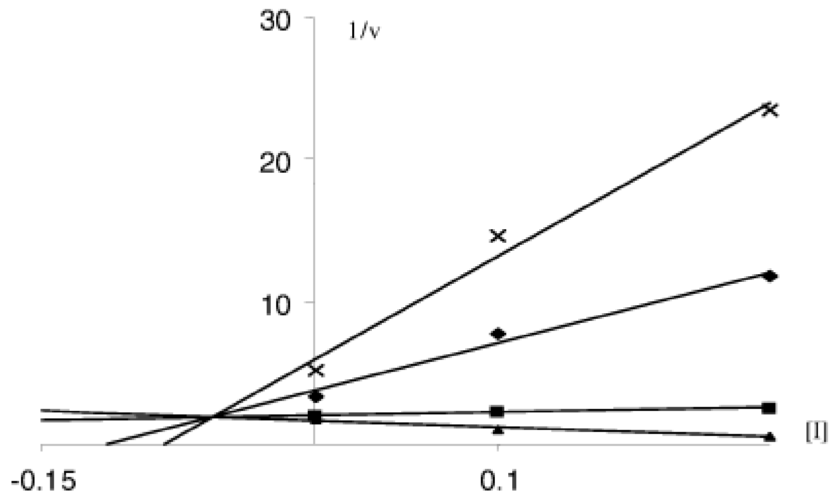


Figure 2.4: Dixon Plot of hydrolysis of sialoside 2.2 by sialidase of *Vibrio cholerae* in the presence of different concentrations of valency corrected polymeric melibiose (18% loading). ×, 0.025; ◆,0.050; ■,0.250; and ▲,2.00 mmol (161).

2.4 Lectin Domain Studies

The roles of carbohydrate binding modules (CBMs) covalently linked to catalytic enzymes have yet to be completely unraveled. New enzymatic studies reveal that the sialidase from *Micromonospora viridifaciens,* which includes a CBM domain, is another in a growing group of modular bacterial enzymes that hydrolyze polymeric substrates much more efficiently than monomeric substrates. Conversely, sialidases lacking these secondary domains, including mammalian sialidases, are not selective toward either substrate form. This knowledge led to the design of our second highly selective polyvalent inhibitor of three modular bacterial sialidases. We also utilize a newly developed HPAEC assay and confirm with SPR studies the lack of selectivity in the lectin domains which allow multivalent interactions between the enzyme and substrate to occur throughout hydrolysis of the surface. The new protocal assisted in the study of steady state kinetics including inhibition using a nonchemical enzymatic assay combined with HPAEC analysis. It is a one pot assay using natural or derivatized substrates on or off of scaffolds and with or without inhibitors requiring no post enzymatic chemical reactions. The assay sensitivity varies with the ability of the detector but is usually sensitive down to low to midnanomolar concentrations.

2.5 Results of the Lectin Domain and *Mv* studies

Here we report the continuation of our findings on multivalency and modular bacterial sialidases focusing on the addition of *Micromonospora viridifaciens* to our growing list of modular enzymes employing the use of multivalency to increase catalytic efficiency and the recognition specificity of the lectin domains involved in these enzymes. We also report a new polyvalent sialic acid based inhibitor that selectively inhibits enzymes containing modular carbohydrate binding domains. The catalytic efficiency of the modular bacterial sialidase *Micromonospora viridifaciens* was tested versus mono- and polyvalent sialyl *N*-acetyl lactosamine substrates (Figure 2.1). *Micromonospora viridifaciens* hydrolyzed monovalent substrate 2.1 with a Michaelis Menten constant 100 times greater than the hydrolysis of polyvalent substrate 2.2 (Figure 2.5), thus agreeing with the previous modular neuraminidases tested (*Clostridium perfringens* and *Vibrio cholerae*) (161). New crystallization data fitting with our previously published mechanistic view have been recently published with galactose in the

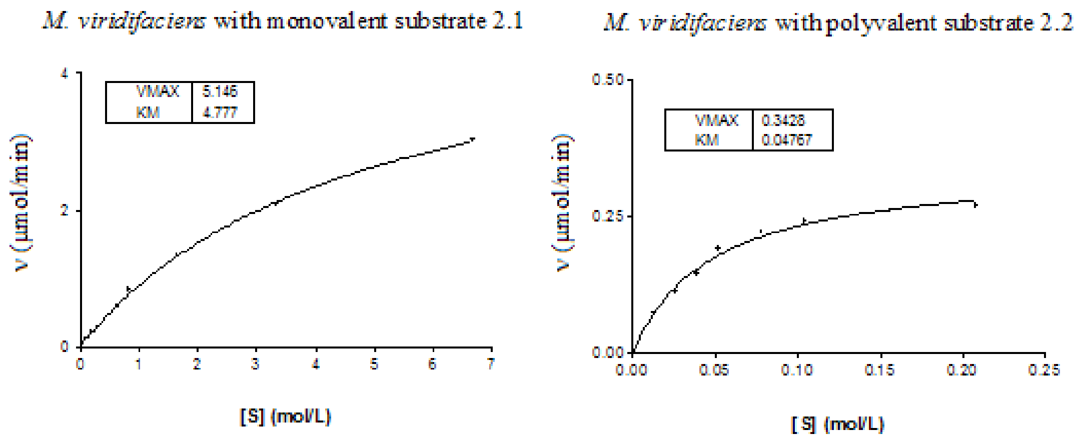


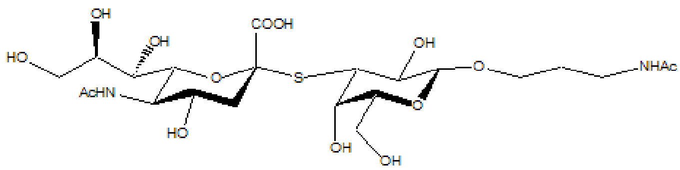
Figure 2.5: Michaelis-Menten curves of *Mv* with monovalent substrate 2.1 and polyvalent substrate 2.2.

carbohydrate binding domain (CBD) of *M. viridifaciens* (180) and sialic acid binding in the CBD of *V. cholerae* (181). This data prompted the further study of the lectin domains themselves concentrating on the specificity of recognized sugars. Using the newly designed HPLC assay, Ic 50 data was determined for the three modular neuraminidases against a varied group of monosaccharides and disaccharides with the non-modular *Salmonella typhimurium* tested as the control enzyme. Table 2.4 shows that each of the modular enzymes have lectins with broad specificities which is a common occurrence with CBDs (182) and may allow the enhanced catalytic ability with a variety of sialosides presented in multivalent arrays such as cell surfaces.

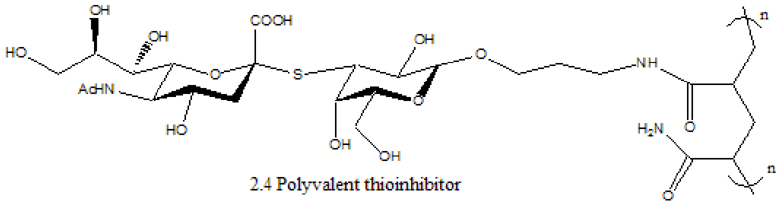
Table 2.4: Ic50 data (in mM) for a variety of sialidases with a constant concentration of 0.125 mM of compound 2.2.

	Vibrio cholerae	*Clostridium perfringens*	*Salmonella typhimurium*	*Micromonospora viridifaciens*
Galactose	16.91	16.41	No inhibition	15.78
Glucose	15.47	21.36	No inhibition	18.80
Fucose	16.77	20.44	No inhibition	No inhibition
Melibiose	18.34	17.88	No inhibition	17.16
Inositol	No inhibition	No inhibition	No inhibition	No inhibition
Ribose	> 30	No inhibition	No inhibition	No inhibition

We believed that a design utilizing both the catalytic and lectin domains to further increase inhibitory effects via multivalency could propel our inhibitor to a potency not seen before with bacterial sialidases. We employed a thio-linked sialyl galactose disaccharide moiety to test this hypothesis (Figure 2.6) (183). This molecule includes a terminal sialic acid but was found to not be recognized by the sialidase catalytic cite of the studied enzymes. The lack of inhibition at the catalytic site was initially attributed to conformational differences between the O- and S-linked glycosidic linkages dictated by the exo-anomeric effect (184). However, further studies of substrate recognition proved that the trisaccharide was necessary for enzyme recognition, as a sialyl-galactose disaccharide substrate failed to produce hydrolyzed sialic acid. Due to the HPLC assay analysis requiring the quantification of freed sialic acid, the ability to test sialic acid as a



2.3 Monovalent thioinhibitor



2.4 Polyvalent thioinhibitor

Figure 2.6: Monovalent (compound 2.3) and polyvalent (2.4) thio-linked Sialyl Galactose

lectin inhibitor had not been possible. This thioglycoside, although not a catalytic inhibitor as per our design, could be used in lectin inhibition assays much as was done with the D-Galactose previously (161).

Inhibition assays of both the monovalent and polyvalent thio-sialyl galactose with *Vibrio cholerae* showed a significant increase in inhibition (Table 2.5) by the polyvalent inhibitor (Ki 19uM) as compared to the monovalent counterpart (Ki 3.2mM) normalized per inhibitor unit (Figure 2.7). The lack of inhibition with *Salmonell typhimurium* proves that all inhibition occurs at the lectin site, thus disallowing multiple interactions to take place. With inhibitor complexation occurring at the lectin domain only, the utility of this type of inhibitor is enhanced due to all known sialidases with modular lectin domains being non-mammalian (185) and thus selectivity towards pathogens has been acheived.

Table 2.5: *Vibrio cholerae* hydrolysis of compounds 2.1 and 2.2 with thio-inhibitors

Substrate	Inhibitor	[I] mM	K_m mM	rel V_{max}	V_{max}/K_m
2.1	None	NA	5.7 ± 0.4	1 ± 0.09	0.18 ± 0.02
2.1	2.3	3	5.1 ± 0.4	0.80 ± 0.06	0.16 ± 0.02
2.2	None	NA	0.04 ± 0.01	0.23 ± 0.02	5.8 ± 1.5
2.2	2.3	1	0.08 ± 0.02	0.34 ± 0.06	4.3 ± 1.1
2.2	2.3	3	0.11 ± 0.02	0.45 ± 0.05	4.1 ± 0.7
2.2	2.3	5	0.19 ± 0.05	0.51 ± 0.04	2.7 ± 0.7
2.2	2.4	0.1	0.29 ± 0.05	0.20 ± 0.02	0.69 ± 0.13
2.2	2.4	0.25	1.1 ± 0.2	0.51 ± 0.04	0.46 ± 0.09
2.2	2.4	0.5	2.7 ± 0.3	0.96 ± 0.09	0.36 ± 0.05

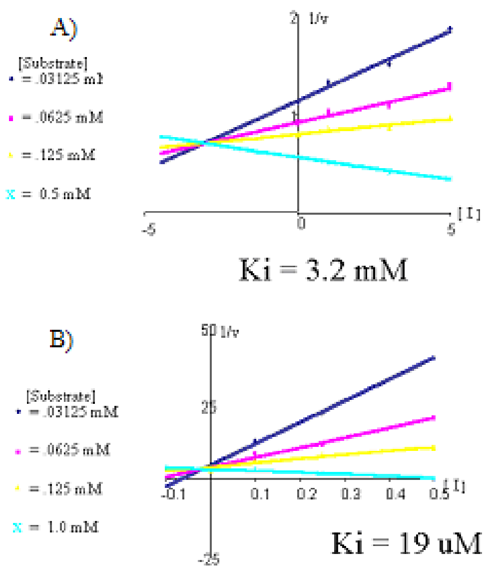


Figure 2.7: Dixon plots of poly-substrate 2.2 with thio-inhibitors **A)** 2.3 (mono) and **B)** 2.4 (poly)

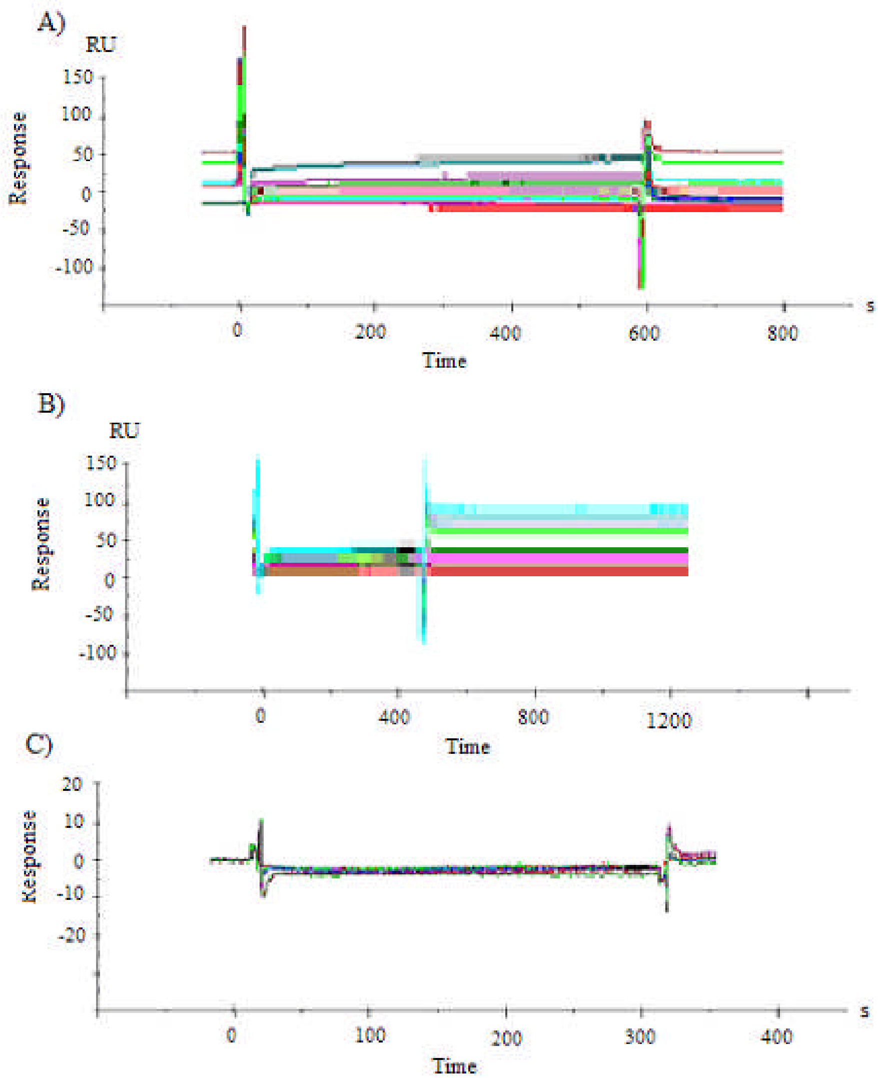


Figure 2.8: Sensorgrams representing the concentration dependent binding of with **A)** sialyl acetylglucsoamine, **B)** melibiose, and **C)** Ribo-inositol with immobilized *Vc* sialidase (7,000 RU). Concentrations from bottom to top of 0.5, 1, 5, 10 and 15 mM were injected over the surface.

Biacore studies were included to test real time binding of a variety of sugars to the lectin domains. Although kinetics for the complexed sugar-lectin interactions could not be determined due to weak affinities, yes/no binding verified binding to certain sugars.

Melibiose, tetracellatraose, and sialylacetylglucosamine all showed binding at low mM concentrations reproducing similar effects as was seen in the enzymatic Ic 50 studies. Ribo-inositol did not show binding at up to 4x these concentrations showing some recognition selectivity within this domain. Oligosaccharide units were used as to increase analyte size for better Biacore measurements (Figure 2.8).

2.6 Experimental procedures

Materials - C. perfringens, V. cholerae, and *S. typhimurium* were purchased from Sigma. Enzymes were diluted in the solutions they were stored in upon arrival. *M. viridifaciens* was given to us by J.Watson. All sugars used in the IC50 experiments were purchased from Sigma.

2.6.1 One pot derivatization-free enzymatic assay

Sialosides 2.1 and 2.2 (25μL) at varying concentrations diluted in an appropriate buffer were incubated at 37^{o}C for 5 minutes. Aliquots of sialidase were added and after incubation of 5 minutes at 37^{o}C, the reactions were quenched by submerging samples in boiling water for 2 minutes. Termination studies were conducted at 15 seconds, 30 seconds, 1 minute, 2 minutes, and 5 minutes followed by cooling back to 37^{o}C overnight. All data shows enzyme death by 15 seconds. All transformations were performed in duplicate. *Vibrio cholerae*, *Micromonospora viridifaciens* and *Salmonella typhimurium* sialidase assays were performed in sodium acetate buffer (pH 5.50, 50mM sodium acetate, 10mM calcium chloride, 50mM sodium chloride in milli-Q water). *Clostridium perfringens* assays were performed in potassium phosphate buffer (pH 5.16, 0.05M potassium phosphate/sodium hydroxide in milli-Q water). A pre-established amount of enzyme was employed to assure that consumption of substrate was less than 15% for initial rate conditions.

2.6.2 Ic50

To a concentration of .125 mM polyvalent 3'-sialyl *N*-acetyllactosamine in an appropriate buffer was added an incrementally increased amount of inhibitor (between 1 and 30 mM) and incubated at 37^{o}C for 5 minutes. To each inhibitor concentration, aliquots of sialidase were added and after incubation of 5 minutes at 37^{o}C, the reactions were quenched by boiling (2 min) followed by cooling on ice. All transformations were performed in duplicate.

2.6.3 Novel HPAEC Determination of Sialic Acid Cleavage

All sugar standards and HPLC eluents were made via the procedure of Hardy and Townsend (186). After undergoing enzymatic transformation, all samples were concentrated to dryness using a Savant AES 1000 Environmental SpeedVac at the low drying rate. The dried samples were then dissolved in 30μL of milli-Q water followed by vortexing (15 seconds) and sonication (5 minutes) to ensure homogeneity of the solution. Samples were loaded into a Metrohm Triathlon Spark autosampler A DIONEX carbopac PA10 column was used in conjunction with a Metrohm 709 IC Pump with a gradient setup using Metrohm IC Net 2.2 software of t = 0 min, NaOH = 95%, NaOAc = 5%; t = 5 min, NaOH = 95%, NaOAc = 5%; t = 25 min, NaOH = 82%, NaOAc = 18%; t = 30 min, NaOH = 82%, NaOAc = 18%; t = 32 min, NaOH = 95%, NaOAc = 5% with a cycle of

60 minutes. Free sialic acid detection was made by a Metrohm 817 Bioscan pulsed amperometric detector. Raw peak areas were determined using Metrohm IC Net 2.2 software. Data analysis, Michaelis curves and Lineweaver-Burke plots were performed with GraphPad Prism version 3.0. Standard curves were run with each assay to account for slight changes in eluent batches.

2.6.4 SPR

The research-grade CM5 sensor chip, HBS-EP buffer, amine coupling reagents (*N*-ethyl-*N'*-dimethylaminopropylcarbodiimide, EDC; *N*-hydroxysuccinimide, NHS; and 1M ethanolamine-HCl-NaOH pH 8.5) were purchased from BIAcore AB (Uppsala Sweden). The biosensor analysis was conducted using a BIAcore 3000 SPR instrument and data were evaluated by using BIAevaluation-2000 software.

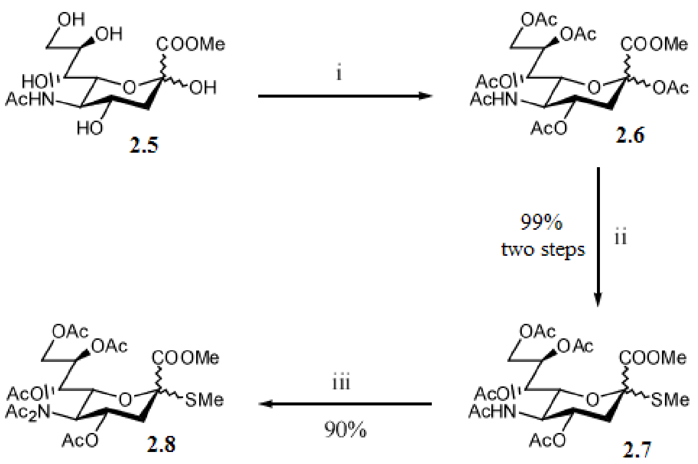
The sialidase from *Vibrio cholerae* was immobilized onto a CM5 research grade sensor chip utilizing the manual technique in HBS-EP running buffer. The surface was activated for 7 minutes using a 1:1 mixture of 100mM NHS and 391mM EDC (both in water). Following activation, the sialidase was diluted to 10μg/ml in 10mM NaOAc pH 4.0. The protein solution was injected over the activated surface for three 5 minute runs for a total of 15 minutes at a flow rate of 10μl/min. The remaining activated surface was blocked by injecting 1.0M ethanolamine-hydrochloride pH 8.5. The final immobilization level was 7000RU. Binding studies were done in PBS buffer pH 7.4.

2.7 Synthetic Routes

2.7.1 Trisaccharide Substrate

In order to study the kinetic behavior of the sialidase enzymes, sufficient quantities of the monovalent trisaccharide 2.1 needed to be synthesized followed by conjugation of a portion of this compound onto the polymer to achieve polyvalent substrate 2.2. Thus, the first stage requires the preparation of the three monosaccharide donors followed by glycosylation in order to prepare final compound 2.1.

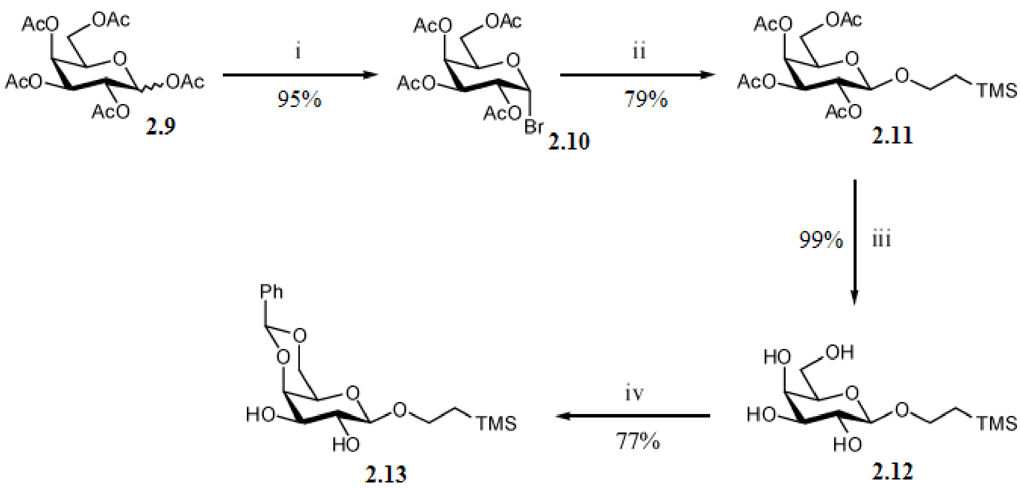
The sialyl donor 2.8 was prepared according to scheme 2.1 from the commercially available starting material 2.5. Per-acetylation of 2.5 was followed by conversion to the thioglycoside 2.7 by treatment with (methylthio)trimethylsilane (TMSSMe) and trimethylsilyl trifluoromethanesulfonic acid (TMSOTf) in 1,2-dichloroethane (DCE). Finally, the sialyl donor 2.8 was finished by diacetylation using *iso*-propenyl acetate and camphosulfonic acid (CSA).



i) Ac_2O, pyridine; ii) TMSSMe, TMSOTf, DCE; iii) $AcOC(CH_3)=CH_2$, TsOH, 60°C

Scheme 2.1: Synthesis of sialyl donor

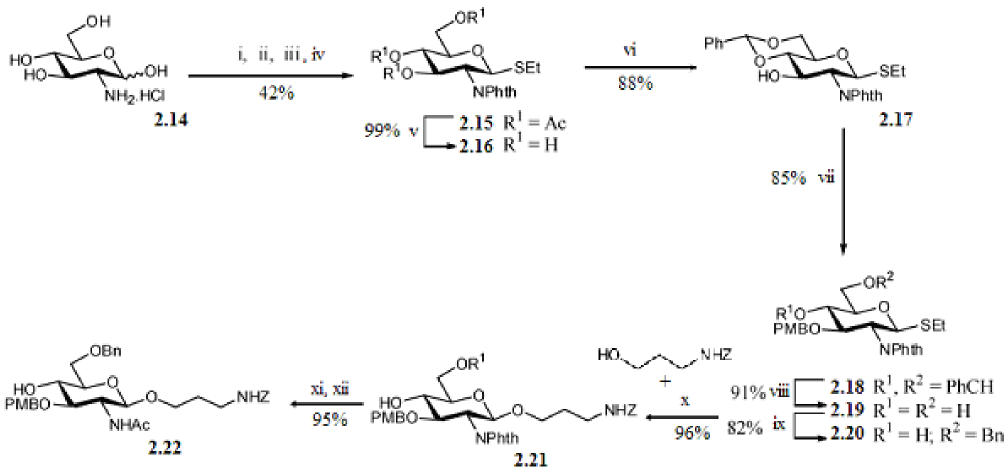
The galactosyl acceptor 2.13 was synthesized following the procedure layed out in scheme 2.2. From the commercially available compound 2.9, bromination via hydrogen bromide in acetic acid gave the bromide 2.10. Subsequent coupling with 2-(trimethylsilyl) ethanol was achieved using a procedure by Hasegawa (187-188). The galactosyl acceptor was finished by deacetylation using sodium methoxide and regioselective benzylidenation at the C-4 and C-6 position.



i) HBr, AcOH, DCM; ii) HgO, $HgBr_2$, $CaSO_4$, TMS-CH_2CH_2OH, DCM; iii) NaOMe, MeOH; iv) $PhCH(OMe)_2$, CSA, MeCN

Scheme 2.2: Synthesis of galactosyl acceptor

The final building block is the *N*-glucosamine acceptor 2.21 with the aminopropyl spacer. The original project for which this trisaccharide was synthesized included the glycosylation of a fucose moiety to finish sialyl lewis x. Although this final target was not necessary for these studies, the abundance of partial building blocks available necessitated a synthetic pathway identical to that prior designed (Scheme 2.3). This final building block started from D-Glucosamine hydrochloride 2.14. Protection via a method described by Kochetkov et al. (189) was followed by synthesizing the thioglycoside 2.15. Deacetylation freed the C-4 and C-6 positions for benzylidene introduction. After protection of C-3 with a *p*-methoxybenzyl group, the benzylidene was deprotected followed by benzylation at C-6. Next the spacer was added to give compound 2.21. Finally, ethylamine mediated cleavage followed by acetylation gave the *N*-glucosamine acceptor 2.22.

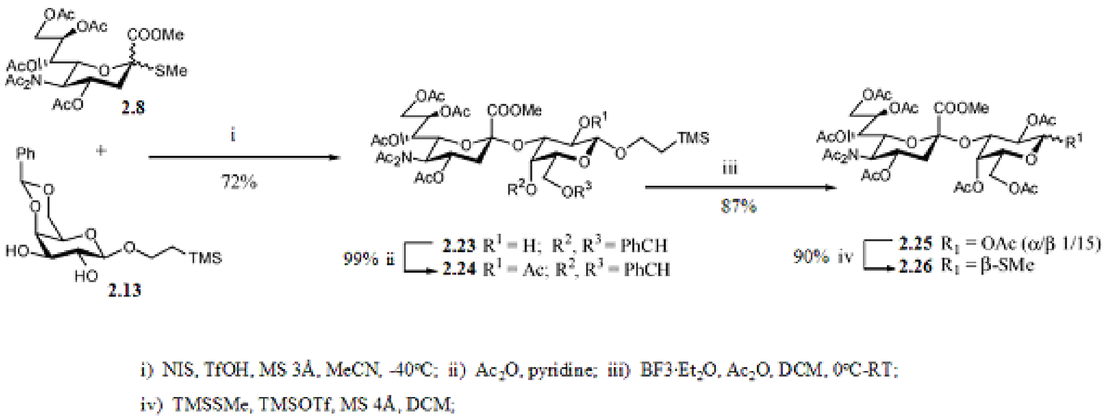


i) Phthalic anhydride, $NaHCO_3$, 40°C; ii) HCl; iii) Ac_2O, NaOAc, reflux; iv) EtSH, $ZnCl_4$, DCM; v) NaOMe, MeOH; vi) $PhCH(OMe)_2$, CSA, MeCN; vii) PMBCl, NaH, TBAI, THF, reflux; viii) 80% AcOH, H_2O, 50°C; ix) $Bu_2Sn(OMe)_2$, toluene, BnBr, TBAI, reflux; x) NIS, TMSOTf, MS 4Å, DCM; xi) EtOH, ethylenediamine, 80°C; xii) MeOH, Ac_2O

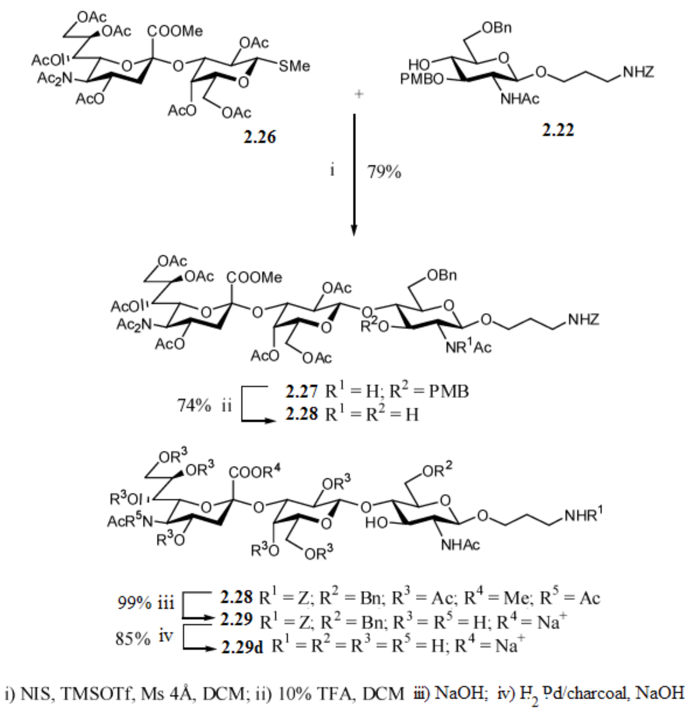
Scheme 2.3: Synthesis of *N*-glucosamine acceptor

Schemes 2.4 and 2.5 show the synthesis of the fully protected trisaccharide. To make the disaccharide donor, glycosylation between sialyl donor 2.8 and galacotsyl acceptor 2.13 was achieved using standard glycosylation conditions for β glycosylation.

Next the galactose moiety was acetylated including O-acetylation at the anomeric center. Conversion of compound 2.25 to a thioglycoside yielded the disaccharide donor. Finally,

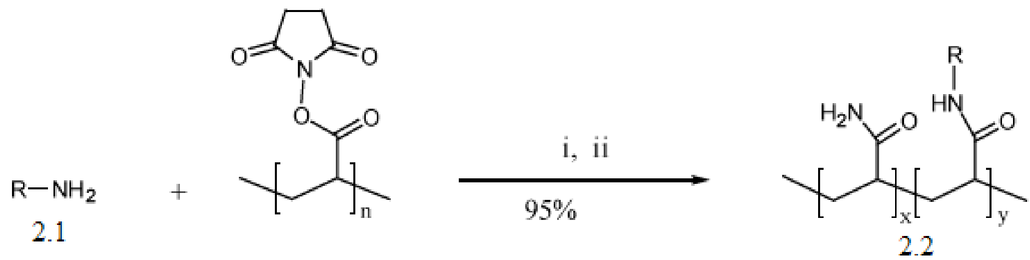


Scheme 2.4: Syntheis of disaccharide donor 2.26



Scheme 2.5: Deprotection of trisaccharide 2.29d

glycosylation of this donor with compound 2.22 gave the fully protected trisaccharide. Deprotection of compound 2.29 (2.29d) followed by acetylation of the linker amine gave monomeric substrate 2.1. Conjugation of the fully deprotected compound gave polymeric substrate 2.2 (Scheme 2.6).

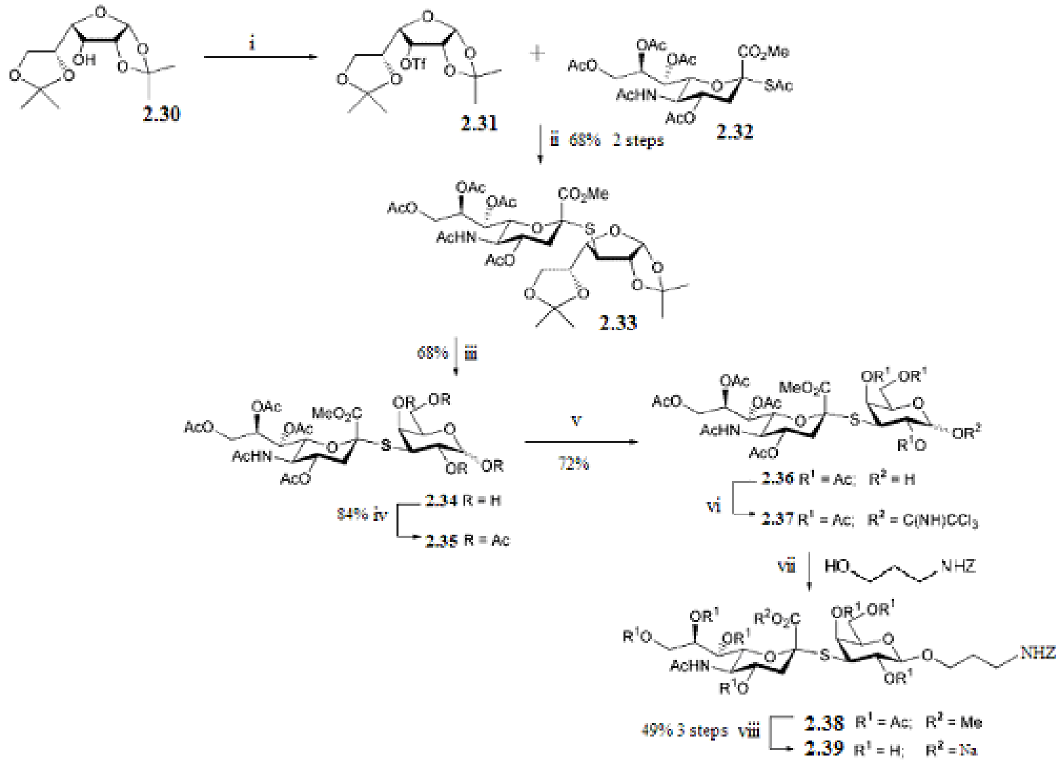


i) DMF, DIPEA; ii) NH_4OH

Scheme 2.6: Polymer conjugation

2.7.2 Thio-Inhibitor

Originally synthesized in hopes of producing a polyvalent inhibitor that binds to both the lectin and catalytic sites of modular sialidases, thio-inhibitors 2.3 and 2.4 allowed us to test sialic acid as a lectin inhibitor as well as provided a second selective inhibitor toward modular bacterial sialidases. From commercially available diacetone glucose, a literature procedure from Lemieux and Stick produces the gulose compound 2.30 at our start (190). Triflation of 1,2,5,6-Di-*O*-isopropylidene-α-D-gulofuranose provides the donor for the glycosylation with the sialyl glycoside 2.32 (191). TFA hydrolysis of the acetals followed by acetic anhydride acetylation gave the acetylated disaccharide 2.33 as 4:1 ratio in the pyranose form. Selective deprotection at the anomeric center was followed by activation as a trichloroacetimidate. Glycosylation with an aminopropyl spacer followed by deprotection gave the thio- disaccharide inhibitor (Scheme 2.7). Acetylation of the spacer amine gave the monomeric inhibitor 2.3 and conjugation onto the polymer gave polymeric inhibitor 2.4 as in scheme 2.5.



i) Tf_2O, pyridine, DCM; ii) Et_2NH, DMF; iii) 90% TFA; iv) Ac_2O, pyridine; v) $N_2H_4\cdot AcOH$, DMF; vi) Cl_3CCN, DBU, DCM; vii) $BF_3\cdot OEt_2$, DCM, 4Å MS; viii) NaOMe, MeOH

Scheme 2.7: Thio-inhibitor synthesis

2.8 Synthetic Data

2.8.1 General

Chemicals for synthesis were purchased from Aldrich, Acros, Sigma or Fluka and used without further purification. Molecular sieves were crushed, activated at 350^{o}C in vacuo for 3 h. Solvents were distilled immediately before use. All reactions were performed under anhydrous conditions unless stated. Monitoring of reactions were done on Kieselgel 60 F_{524} thin-layer chromatography plates (Merck) and visualized under UV light (254 nm) or by charring with 10% sulfuric acid in methanol. Chromatography was performed on silica gel 60 (Merck, 70-230 mesh) and flash chromatography on silica gel 60 (Merck, 63-85 mesh). NMR data were recorded on Varian Inova-300, -500, or –600 spectrometers using Sun workstations. Chemical shifts were measured in parts per million and coupling constants in Hertz. Matix assisted laser desorption ionization – time of flight mass spectra were recorded using a Hewlett Packard G2025A system using *trans*-3-indoleacrylic acid or gentisic acid as matrices.

2.8.2 Compound synthesis

Methyl [methyl 5-(*N*-acetylacetamido)-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-2-thio-D-*glycero*-D-*galacto*-non-2-ulopyranosid]onate (2.8)

A solution of commercial methyl (5-acetamido-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-ulopyranosid)onate **(2.5)** (10.00 g, 30.93 mmol) in acetic anhydride (50 mL) and pyridine

(100 mL) was stirred at room temperature for 18 h followed by concentration *in vacuo* and then co-evaporated with toluene (3 x 5 mL), ethanol (2 x 5 mL) and DCM (2 x 5 mL) to afford **2.6** as a yellow oil in a quantitative yield and used without purification (16.42 g). R_f α = 0.22, β = 0.17 (acetone/toluene, 3/7, v/v); FAB MS *m/z* = 556 $[M + Na]^+$.

To the crude product **2.6** in dry 1,2-dichloroethane (100 mL)**,** TMSSMe (8.8 mL, 61.86 mmol) and TMSOTf (4.2 mL, 23.20 mmol) were added at 50^{o}C and left to stir under an atmosphere of argon. TLC analysis (acetone/toluene, 3/7, v/v) after 18 h indicated completion of the reaction to the product (R_f α = 0.28, β = 0.23). The reaction mixture was neutralized with Et_3N, diluted with DCM (100 mL) and washed with saturated aqueous $NaHCO_3$ (2 x 30 mL) and water (3 x 30 mL). The combined organic layers were dried with $MgSO_4$, filtered and concentrated *in vacuo* to afford **2.7** as a yellow foam containing a mixture of α/β = 3/7 (16.1 g, quantitative yield), which was used without purification. FAB MS *m/z* = 544 $[M + Na]^+$; ^{1}H NMR($CDCl_3$): δ 6.13 - 3.90 (envelope, 8H, α/β: H-4, H-5, H-6, H-7, H-8, H-9a, H-9b, NH), 3.80 (s, 3H, α–$COOCH_3$), 3.79 (s, 3H, β–$COOCH_3$), 2.71 (dd, 1H, H-3eα, $J_{3e\alpha,3e\alpha}$ = 12.5 Hz, $J_{3e\alpha,4\alpha}$ = 4.8 Hz), 2.52 (dd, 1H, H-3eβ, $J_{3e\beta,3e\beta}$ = 14.0 Hz, $J_{3e\beta,4\beta}$ = 5.2 Hz), 2.16 - 1.86 (envelope, 19H, α/β: 3a, 5 x CO-CH_3, SCH_3); ^{13}C NMR($CDCl_3$): δ 171.0, 170.2, 167.9 (α/β: 5 x $\underline{C}$O-CH_3, C-1), 84.7, 82.9 (C-2α, C-2β), 74.0, 72.5, 72.1, 69.8, 69.3, 68.6, 67.3, 56.5, 53.0, 49.5 (α/β: C-4, C-5, C-6, C-7, C-8, COO$\underline{C}H_3$), 62.4, 62.2 (C-9α, C-9β), 37.8, 37.0 (C-3α, C-3β), 23.2, 21.2, 21.1, 20.9 (α/β: 5 x CO-$\underline{C}H_3$), 12.1, 11.4 (α–SCH3, β–SCH3).

A solution of **2.7** (16.1 g, 30.9 mmol) and camphorsulfonic acid (0.72 g, 3.09 mmol) in *iso*-propenyl acetate (100 mL) was stirred for 16 h under an atmosphere of argon at 60^{o}C. The mixture was neutralized with Et_3N, concentrated *in vacuo* and the resulting residue purified by silica gel column chromatography (acetone/toluene, 1/9, v/v) yielding **2.8** as a white foam in an α/β = 3/7 mixture (15.66 g, 90%). R_f α = 0.52, β = 0.47 (acetone/toluene, 3/7, v/v); FAB MS *m/z* = 586 $[M + Na]^+$; ^1H NMR ($CDCl_3$): δ 6.15 - 4.10 (envelope, 7H, α/β: H-4, H-5, H-6, H-7, H-8, H-9a,H-9b), 3.84 (s, 3H, α−COOCH3), 3.80 (s, 3H, β−$COOCH_3$), 2.88 (dd, 1H, H-3eα, $J_{3e\alpha,3e\alpha}$ = 12.5 Hz, $J_{3e\alpha,4\alpha}$ = 5.1 Hz), 2.68 (dd, 1H, H-3eβ, $J_{3e\beta,3e\beta}$ = 13.6 Hz, $J_{3e\beta,4\beta}$ = 5.1 Hz), 2.15 - 1.97 (m, 22H, α/β: 3a, 6 x CO-CH_3, SCH_3); ^{13}C NMR ($CDCl_3$): δ 174.3, 173.4, 170.4, 170.1, 169.4, 168.6, 167.5, (α/β: 6 x <u>C</u>O-CH_3, <u>C</u>OOCH_3), 84.7, 82.6 (C-2α, C-2β), 71.5, 71.3, 68.9, 68.6, 68.3, 68.2, 67.2, 67.1, 57.3, 56.9, 52.9, 52.7 (α/β: C-4, C-5, C-6, C-7, C-8, COO<u>C</u>H_3), 61.9, 61.7 (C-9α, C-9β), 38.8, 38.1 (C-3α, C-3β), 27.9, 25.9, 21.3, 21.0, 20.8, 20.6 (α/β: 6 x CO-<u>C</u>H_3), 12.0, 11.5 (α−SCH3, β−SCH3).

2-(Trimethylsilyl) ethyl 2,3,4,6-tetra-*O*-acetyl-β−D-galactopyranoside (2.11)

To a solution of 1,2,3,4,6-penta-*O*-acetyl-D-galactopyranose (**2.9)** (20.00 g, 51.24 mmol) in dry DCM (80 mL) was added a 30% solution of HBr in acetic acid (80 mL). After stirring for 4 h at room temperature under an atmosphere of nitrogen, the mixture was diluted with DCM (80 mL), poured into ice-water (200 mL) and the organic layer washed with saturated $NaHCO_3$ (3 x 50 mL), saturated NaCl (3 x 50 mL) and water (3 x 50 mL). The combined filtrates were dried ($MgSO_4$), filtered and the solvent evaporated *in vacuo* to afford compound **2.10** as a pale yellow syrup and used without further purification. (20.08 g, 95%); R_f 0.62 (acetone/DCM, 1/19, v/v); MALDI TOF *m/z* = 434 $[M + Na]^+$. To a mixture of compound **2.10** (20.08 g, 49.75 mmol), mercuric(II) oxide (10.78 g, 49.75 mmol), mercuric(II) bromide (0.90 g, 2.49 mmol) and $CaSO_4$ (13.56 g, 99.71

mmol) was added a solution of 2-(trimethylsilyl)ethanol (10.7 mL, 74.78 mmol) in dry DCM (160 mL). After stirring for 90 h at room temperature, TLC analysis (acetone/DCM, 5/95, v/v) indicated the conversion of **2.10** (R_f 0.62) to the product **2.11** (R_f 0.49). The reaction mixture was diluted with DCM (60 mL), filtered and washed with saturated aqueous $NaHCO_3$ (3 x 30 mL) and water (3 x 30 mL). The combined organic extracts were dried ($MgSO_4$), filtered, concentrated under reduced pressure and the residue purified by silica gel column chromatography (eluent: DCM-DCM/acetone, 95/5, v/v). Compound **2.11** was isolated as a pale yellow oil in 79% yield (17.98 g). FAB MS *m/z* = 471 $[M + Na]^+$; ^{1}H NMR ($CDCl_3$): δ 5.37 (dd, 1H, H-4, $J_{4,5}$ = 1.1 Hz, $J_{3,4}$ = 3.3 Hz), 5.19 (dd, 1H, H-2, $J_{2,3}$ = 10.7 Hz, $J_{1,2}$ = 8.1 Hz), 5.00 (dd, 1H, H-3), 4.47 (d, 1H, H-1), 4.20 (dd, 1H, H-6a, $J_{6a,b}$ = 11.2 Hz, $J_{5,6a}$ = 6.6 Hz), 4.11 (dd, 1H, H-6b, $J_{5,6b}$ = 7.0 Hz), 4.02 - 3.93 (m, 1H, O-CH_{2a}), 3.89 (dt, 1H, H-5), 3.60 - 3.51 (m, 1H, O-CH_{2b}), 2.14, 2.04, 2.03, 1.97 (4s, 12H, 4 x CO-CH_3), 1.03 - 0.95 (m, 2H, O-$CH_2C\underline{H}_2$ 0.00 [s, 9H, $Si(CH_3)_3$]; ^{13}C NMR(CD_3OD): δ 172.5, 172.3, 172.9, 171.8 (4 x $\underline{C}O$-CH_3), 102.1 (C-1). 73.1, 72.2, 71.0, 69.4 (C-2, C-3, C-4, C-5), 68.8 (C-6), 63.1 (O-CH2), 21.3 - 21.0 (4 x CO-$\underline{C}H_3$), 19.3 (O-$CH_2\underline{C}H_2$), -0.8 [Si(CH3)3].

2-(Trimethylsilyl) ethyl 4,6-*O*-benzylidene-β– D-galactopyranoside (2.13)

A solution of compound **2.11** (17.9 g, 39.86 mmol), sodium methoxide (4.50 g, 83.33 mmol) and methanol (100 mL) was stirred at room temperature for 2 h. The reaction mixture was neutralized with DOWEX-50 (H^+) ion exchange resin, filtered, concentrated *in vacuo* and the residue co-evaporated from toluene (2 x 5 mL). The deprotected compound **2.12** was obtained in a quantitative yield (11.20 g) and used directly in the next step. Benzaldehyde dimethyl acetal (12.4 mL, 82.72 mmol) was added to a solution of compound **2.12** (11.60 g, 41.36 mmol) in dry acetonitrile (160 mL) and the solution acidified with camphorsulfonic acid (0.48 g, 2.07 mmol) to pH = 3 and stirred at room temperature under an atmosphere of argon for 18 h. TLC analysis (acetone/toluene, 3/7, v/v) indicated full conversion of the starting material **2.12** (R_f 0.00) to compound **2.13** (R_f 0.22). The mixture was neutralized by Et_3N and concentrated *in vacuo* to dryness. Purification of the residue by silica gel column chromatography (eluent: toluene-toluene/acetone, 7/3, v/v) gave compound **2.13** as an amorphous white solid (11.48 g, 77%). FAB MS *m/z* = 391 $[M + Na]^+$; ^{1}H NMR($CDCl_3$): δ 7.49 - 7.45 (m, 2H, $C_6\underline{H}_5CH$), 7.38 - 7.31 (m, 3H, $C_6\underline{H}_5CH$), 5.53 (s, 1H, $C_6H_5C\underline{H}$), 4.33 (dd, 1H, H-6a, $J_{5,6a}$ = 1.8 Hz, $J_{6a,6b}$ = 12.5 Hz), 4.27 (d, 1H, H-1, $J_{1,2}$ = 7.4 Hz), 4.19 (dd, 1H, H-4, $J_{4,5}$ = 1.1 Hz, $J_{3,4}$ = 3.3 Hz), 4.06 (dd, 1H, H-6b, $J_{5,6b}$ = 1.8 Hz), 4.09 - 4.00 (m, 1H, O-CH_{2a}), 3.72 (t, 1H, H-2), 3.65 (dd, 1H, H-3, $J_{2,3}$ = 9.6 Hz), 3.59 - 3.50 (m, 1H, O-CH_{2b}), 3.47 - 3.45 (dd, 1H, H-5), 2.49, 2.53 (2s, 2H, 3-OH, 2-OH), 1.04 - 0.97 (m, 2H, O-$CH_2C\underline{H}_2$), 0.00 [s, 9H, $Si(CH_3)_3$; ^{13}C NMR ($CDCl_3$): δ 137.6 (Cq, $\underline{C}_6H_5CH$), 129.2, 128.2 (x 2), 126.5 (x 2) (5 x $\underline{C}_6H_5CH$), 102.3, 101.4 ($C_6H_5C\underline{H}$, C-1), 75.4, 72.7, 71.7, 66.6 (C-2, C-3, C-4, C-5), 69.2 (C-6), 67.4 (O-CH2), 18.3 (O-$CH_2\underline{C}H_2$), -1.33 [Si(CH3)$_3$].

Ethyl 3-*O*-(*p*-methoxybenzyl)-4,6-*O*-benzylidene-2-deoxy-2-phthalimido-1-thio-β–D–glucopyranoside (2.18)

Ethyl 4,6-*O*-benzylidene-2-deoxy-2-phthalimido-1-thio-β–D-glucopyranoside (**2.17)** (192) (3.50 g, 7.93 mmol) was dissolved in dry THF (140 mL) followed by the addition of sodium hydride (60%) suspended in mineral oil (0.48 g, 11.89 mmol). After stirring the suspension for 10 minutes *p*-methoxybenzyl chloride (1.61 mL, 11.89 mmol) and tetrabutylammonium iodide (TBAI) (2.93 g, 7.93 mmol) were added to the reaction mixture heated to reflux. After a period of 18 h, TLC analysis (ethyl acetate/hexane, 2/5, v/v) indicated conversion of **2.17** (R_f 0.16) to the product **2.18** (R_f 0.29). The reaction mixture was diluted with diethyl ether (50 mL), poured into ice-water (100 mL) and the organic layer washed with saturated NH_4Cl (3 x 30 mL), and water (3 x 30 mL). The combined organic extracts were dried ($MgSO_4$), filtered and concentrated *in vacuo.* The residual oil was purified by silica gel column chromatography (eluent gradient: hexane to hexane/ethyl acetate, 7/3, v/v) to afford **2.18** as a pale yellow foam in 91% yield (4.06 g). MALDI TOF *m/z* = 584 $[M + Na]^+$; ^1H NMR($CDCl_3$): δ 7.84 - 7.82 (m, 1H, Phth), 7.75 - 7.62 (m, 3H, Phth), 7.55 - 7.51 (m, 2H, $C_6\underline{H}_5CH$), 7.44 - 7.36 (m, 3H, $C_6\underline{H}_5CH$), 6.92 - 6.90 (d, 2H, $OCH_3C_6\underline{H}_4CH_2$, $J_{CH2A,CH2B}$ = 8.5 Hz), 6.39 - 6.36 (d, 2H, $OCH_3C_6\underline{H}_4CH_2$), 5.63 (s, 1H, $C_6H_5C\underline{H}$), 5.34 (d, 1H, H-1, $J_{1,2}$ = 10.7 Hz), 4.71 (d, 1H, $OCH_3C_6H_4C\underline{H}_{2a}$, $J_{CH2a,CH2b}$ = 12.3 Hz), 4.44 – 4.39 (m, 3H, H-3, $OCH_3C_6H_4C\underline{H}_{2b}$, H-6a), 4.26 (t, 1H, H-2, $J_{2,3}$ = 10.0 Hz), 3.87 - 3.77 (m, 2H, H-6b, H-4), 3.73 - 3.69 (m, 1H, H-5), 3.62 (s, 3H, OCH_3), 2.71 - 2.55 (m, 2H, S-CH_2), 1.16 (t, 3H, S-$CH_2C\underline{H}_3$); ^{13}C NMR ($CDCl_3$): δ 167.2, 167.1 (2 x NCO), 159.1 (Cq, $OCH_3\underline{C}_6H_4CH_2$), 137.6 (Cq, $\underline{C}_6H_5CH$), 134.0 (2 x Phth), 131.5, 131.3 (2 x Cq, Phth), 130.2 (Cq, OCH3C6H4CH2), 130.0 (2 x $OCH_3\underline{C}_6H_4CH_2$), 129.2, 128.5, 126.3 (5 x $\underline{C}_6H_5CH$), 123.7, 123.3 (2 x Phth), 113.6 (2 x $OCH_3\underline{C}_6H_4CH_2$), 101.6 ($C_6H_5\underline{C}H$), 83.3 (C-4), 82.0 (C-1), 75.3 (C-3), 74.0 ($OCH_3C_6H_4\underline{C}H_2$), 70.1 (C-5) 69.0 (C-6), 55.9 (C-2), 55.1 (OCH3), 24.3 (S-CH2), 15.1 (S-$CH_2\underline{C}H_3$).

Ethyl 3-*O*-(*p*-methoxybenzyl)-2-deoxy-2-phthalimido-1-thio-β– D-glucopyranoside (2.19)

Compound **2.18** (4.06 g, 7.23 mmol) was added to a solution of 80% acetic acid in water and heated to 50^{o}C. After 4 h, the reaction mixture was cooled, concentrated *in vacuo* and then co-evaporated with toluene (3 x 5 mL). The residual oil was dissolved in DCM (100 mL) and washed with saturated aqueous $NaHCO_3$ (3 x 25 mL) and water (3 x 25 mL). The combined organic extracts were dried ($MgSO_4$), filtered, concentrated under reduced pressure and purified by silica gel column chromatography (eluent gradient: toluene to toluene/acetone, 7/3, v/v). Compound **2.19** was obtained as a white syrup in 91% yield (3.11 g). R_f 0.19 (toluene/acetone, 7/3, v/v); MALDI TOF *m/z* = 495 [M + Na]$^+$; ^{1}H NMR ($CDCl_3$): δ 7.83 - 7.82 (m, 1H, Phth), 7.71 (bs, 3H, Phth), 7.00 (d, 2H, $OCH_3C6\underline{H}_4CH_2$, $J_{CH2A,CH2B}$ = 8.6 Hz), 6.53 (d, 2H, $OCH_3C_6\underline{H}_4CH_2$), 5.31 (d, 1H, H-1, $J_{1,2}$ = 9.6 Hz), 4.59 (d, 1H, $OCH_3C_6H_4C\underline{H}_{2a}$, $J_{CH2aCH2b}$ = 12.3 Hz), 4.47 (d, 1H, $OCH_3C_6H_4C\underline{H}_{2b}$), 4.29 (dd, 1H, H-3, $J_{2,3}$ = 9.9 Hz, $J_{3,4}$ = 8.4 Hz), 4.20 (t, 1H, H-2), 3.95 (dd, 1H, H-6a, $J_{5,6a}$ = 3.6 Hz, $J_{6a,6b}$ = 11.7 Hz), 3.84 (dd, 1H, H-6b, $J_{5,6b}$ = 4.8 Hz), 3.75 (dd, 1H, H-4, $J_{4,5}$ = 9.6 Hz), 3.65 (s, 3H, OCH_3), 3.60 – 3.54 (m, 1H, H-5), 2.67 - 2.59 (m, 2H, S-CH_2), 1.17 (t, 3H, S-$CH_2C\underline{H}_3$); ^{13}C NMR ($CDCl_3$): δ 168.3, 167.6 (2 x NCO), 159.1 (Cq, $OCH_3\underline{C}_6H_4CH_2$), 134.1 (2 x Phth), 131.8, (2 x Cq, Phth), 130.3 (Cq, $OCH_3\underline{C}_6H_4CH_2$), 129.7 (2 x $OCH_3\underline{C}_6H_4CH_2$), 123.7, 123.4 (2 x Phth), 113.8 (2 x $OCH_3\underline{C}_6H_4CH_2$), 81.6 (C-1), 80.2 (C-3), 79.7 (C-5), 74.5 ($OCH_3C_6H_4\underline{C}H_2$), 72.5 (C-4), 62.9 (C-6), 55.2 (OCH_3), 55.0 (C-2), 24.6 (S-CH_2), 15.3 (S-$CH_2\underline{C}H_3$).

Ethyl 3-*O*-(*p*-methoxybenzyl)-6-*O*-benzyl-2-deoxy-2-phthalimido-1-thio-β– D-glucopyranoside (2.20)

To compound **2.19** (1.62 g, 3.42 mmol) in dry toluene (100 mL) was added dibutyl tin dimethoxide (0.86 mL, 3.76 mmol). The reaction mixture was refluxed under Dean-Stark conditions for 3h, after which half of the solvent was removed. Once the solution reached room temperature, BnBr (0.61 mL, 5.13 mmol) and TBAI (1.89 g, 5.13 mmol) were added and the reaction mixture refluxed for a further for 2h. The solution was diluted with toluene (50 mL) and the filtrate washed with aqueous 1M KF (3 x 15 mL), saturated aqueous $NaHCO_3$ (3 x 15 mL) and water (3 x 15 mL). The combined organic extracts were dried ($MgSO_4$), filtered and concentrated *in vacuo* and the residue was purified by silica gel column chromatography (eluent gradient: toluene to toluene/acetone, 7/3, v/v) to obtain compound **2.20** as an oil (1.58g, 82%). R_f 0.59 (toluene/acetone, 7/3, v/v); MALDI TOF *m/z* = 586 $[M + Na]^+$; ^1H NMR($CDCl_3$): δ 7.82 - 7.80 (m, 1H, Phth), 7.71 - 7.68 (m, 3H, Phth), 7.36 - 7.28 (m, 5H, $C_6\underline{H}_5CH_2$), 6.96 (d, 2H, $OCH_3C_6\underline{H}_4CH_2$, $J_{CH2A,CH2B}$ = 8.5 Hz), 6.46 (d, 2H, $OCH_3C_6\underline{H}_4CH_2$), 5.27 (d, 1H, H-1, $J_{1,2}$ = 9.6 Hz), 4.66 (d, 1H, $OCH_3C_6H_4C\underline{H}_{2a}$, $J_{CH2a,CH2b}$ = 12.1 Hz), 4.66 – 4.56 (m, 2H, $C_6H_5C\underline{H}_2$), 4.48 (d, 1H, $OCH_3C_6H_4C\underline{H}_{2b}$), 4.24 (dd, 1H, H-3, $J_{2,3}$ = 10.2 Hz, $J_{3,4}$ = 8.0 Hz), 4.19 (dd, 1H, H-2), 3.85 (dd, 1H, H-6a, $J_{5,6a}$ = 4.9 Hz, $J_{6a,6b}$ = 9.8 Hz), 3.84 – 3.75 (m, 2H, H-4, H-6b), 3.69 – 3.65 (m, 1H, H-5), 3.62 (s, 3H, OCH_3), 2.96 (d, 1H, 4-OH, $J_{4,OH}$ = 1.8 Hz), 2.70 - 2.50 (m, 2H, S-CH_2), 1.15 (t, 3H, S-$CH_2C\underline{H}_3$); ^{13}C NMR ($CDCl_3$): δ 168.4 (2 x NCO), 159.2 (2 x Cq, $OCH_3\underline{C}_6H_4CH_2$), 137.9 (Cq, $\underline{C}_6H_5CH_2$), 134.0 (2 x Phth), 131.8 (2 x Cq, Phth), 129.8 (2 x $OCH_3\underline{C}_6H_4CH_2$), 129.0 – 128.0 (5 x $\underline{C}_6H_5CH_2$), 123.6, 123.4 (2 x Phth), 113.8 (2 x $OCH_3\underline{C}_6H_4CH_2$), 81.4 (C-1), 79.5 (C-3), 77.9 (C-5), 74.7 (C-4), 74.2 ($OCH_3C_6H_4\underline{C}H_2$), 73.9 ($C_6H_5\underline{C}H_2$), 71.1 (C-6), 55.1 ($OCH_3$), 54.7 (C-2), 24.1 (S-$CH_2$), 15.0 (S-$CH_2\underline{C}H_3$).

3-(*N*-Benzoyloxycarbonyl) aminopropyl 3-*O*-(*p*-methoxybenzyl)-6-*O*-benzyl-2-deoxy-2-phthalimido-β– D-glucopyranoside (2.21)

A mixture of 3-(benzyloxycarbonylamino)-1-propanol (516 mg, 2.47 mmol) and compound **2.20** (744 mg, 1.37 mmol) with activated crushed molecular sieves (1.00 g, 4 Å) in DCM (30 mL) was stirred for 3 h under an atmosphere of argon. NIS (463 mg, 2.06 mmol) and TMSOTf (20 μL, 0.11 mmol) were added. After stirring the reaction mixture for 10 min at room temperature, the reaction was diluted with DCM (25 mL), filtered, and the residue washed with DCM (3 x 10 mL). The combined filtrates were washed with 15% aqueous $Na_2S_2O_3$ (15 mL, w/v), saturated aqueous $NaHCO_3$ (2 x 15 mL) and H2O (2 x 15 mL). The organic phase was dried ($MgSO_4$) and filtered, and the filtrate was concentrated *in vacuo.* Purification by silica gel column chromatography (eluent gradient: toluene to toluene/ethyl acetate, 1/1, v/v) afforded **2.21** as a white foam (0.98 mg, 98%). R_f 0.44 (acetone/toluene, 3/7, v/v); MALDI TOF *m/z* = 733 $[M + Na]^+$; ^1H NMR ($CDCl_3$): δ 7.66 - 7.64 (m, 4H, Phth), 7.38 - 7.26 (m, 10H, $C_6\underline{H}_5CH_2$, $C_6\underline{H}_5CH_2OCO$), 6.95 (d, 2H, $OCH_3C_6\underline{H}_4CH_2$, $J_{CHA,CHB}$ = 8.5 Hz), 6.46 (d, 2H, $OCH_3C_6\underline{H}_4CH_2$), 5.16 (d, 1H, H-1, $J_{1,2}$ = 8.0 Hz), 5.14 - 5.08 (m, 1H, NH), 5.01 (s, 2H, $C_6H_5C\underline{H}_2OCO$), 4.64 (d, 1H, $OCH_3C_6H_4C\underline{H}_{2a}$, $J_{CH2aCH2b}$ = 12.1 Hz), 4.59 - 4.52 (m, 2H, $C_6H_5C\underline{H}_2$), 4.46 (d, 1H, $OCH_3C_6H_4C\underline{H}_{2b}$), 4.16 (dd, 1H, H-3, $J_{2,3}$ = 10.4 Hz, $J_{3,4}$ = 7.7 Hz) 4.10 (dd, 1H, H-2), 3.81 - 3.70 (m, 4H, O-CH_{2a}, H-4, H-6a,H-6b), 3.67 - 3.60 (m, 1H, H-5), 3.63 (s, 3H, OCH_3), 3.54 - 3.47 (m, 1H, O-CH_{2b}), 3.16 - 3.02 (m, 2H, CH_2-N), 2.78 (s, 1H, 4-OH) 1.68 - 1.55 (m, 2H, O-$CH_2C\underline{H}_2CH_2$); ^{13}C NMR ($CDCl_3$): δ 168.0, 167.5 (2 x NCO), 159.0, (2 x Cq, $OCH_3\underline{C}_6H_4CH_2$), 156.5 (NH-COO), 137.8, (Cq, $\underline{C}_6H_5CH_2$), 136.9 (Cq, $\underline{C}_6H_5CH_2OCO$), 134.0 (2 x Phth), 131.8, 130.5 (2 x Cq, Phth), 129.7 (2 x $OCH_3\underline{C}_6H_4CH_2$), 128.6, 128.2, 128.1, 128.0 (10 x $\underline{C}_6H_5CH_2$, $\underline{C}_6H_5CH_2OCO$), 123.4 (2 x Phth), 113.7 (2 x $OCH_3\underline{C}_6H_4CH_2$), 98.6 (C-1), 78.8 (C-3), 74.2, 74.0, 73.9, 73.8, (C-4, C-5, $C_6H_5\underline{C}H_2$, $OCH_3C_6H_4\underline{C}H_2$), 70.7 (C-6), 67.2 ($C_6H_5\underline{C}H_2OCO$), 66.7 (O-$CH_2$), 55.7 (C-2), 55.2 ($OCH_3$), 38.1 ($CH_2$-N), 29.7 (O-$CH_2\underline{C}H_2CH_2$).

2-(Trimethylsilyl) ethyl *O*-{[methyl 5-(*N*-acetylacetamido)-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-*glycero*-α– D-*galacto*-non-2-ulopyranosid]-onate}-(2-3)-4,6-*O*-benzylidene-β–D-galactopyranoside (2.23)

A suspension of **2.8** (1.76 g, 3.13 mmol) and **2.13** (1.00 g, 2.72 mmol) with activated crushed molecular sieves (2.50 g, 3Å) in dry acetonitrile (30 mL) was stirred at room temperature for 2 h. The mixture was cooled to -40° C and NIS (1.0 g, 5.63 mmol) and TfOH (47 µL, 0.54 mmol) were added. After 10 minutes, TLC analysis (acetone/toluene, 3/7, v/v) showed conversion of the starting materials into product (Rf 0.42). The reaction mixture was diluted with DCM (30 mL), filtered, and washed with aqueous $Na_2S_2O_3$ (15%, 20 mL, w/v), saturated aqueous $NaHCO_3$ (3 x 15 mL) and water (3 x 15 mL). The combined organic extracts were dried ($MgSO_4$), filtered and concentrated *in vacuo.* The residue was purified by silica gel column chromatography (eluent gradient: toluene to toluene/acetone, 7/3, v/v) to give compound **2.23** as a white foam in 72% yield (1.73 g). MALDI TOF *m/z* = 907 $[M + Na]^+$; ^{1}H NMR ($CDCl_3$): δ 7.50 - 7.40 (m, 2H, $C_6\underline{H}_5CH$), 7.35 - 7.25 (m, 3H, $C_6\underline{H}_5CH$), 5.48 (ddd, 1H, H-4’, $J_{4',5'}$ = 10.0 Hz,), 5.40 (s, 1H, $C_6H_5C\underline{H}$), 5.37 (ddd, 1H, H-8’, $J_{7'.8'}$ = 8.5 Hz, $J_{8',9a'}$ = 2.8 Hz, $J_{8',9b'}$ = 5.5 Hz), 5.17 (dd, 1H, H-7’, $J_{6',7'}$ = 2.0 Hz), 4.95 (dd, 1H, H-6’, $J_{5',6'}$ = 10.0 Hz), 4.46 (d, 1H, H-1, $J_{1,2}$ = 7.5 Hz), 4.35 (dd, 1H, H-9a’, $J_{9a',9b'}$ = 12.5 Hz), 4.28 (dd, 1H, H-6a, $J_{5,6a}$ = 2.0 Hz, $J_{6a,6b}$ = 12.0 Hz), 4.26 (dd, 1H, H-3, $J_{2,3}$ = 10.0 Hz, $J_{3,4}$ = 3.5 Hz), 4.17 (ρt, 1H, H-5’), 4.09 (dd, 1H, H-9b’), 4.08 (dd, 1H, H-6b, $J_{5,6b}$ = 2.0 Hz), 4.08 - 4.03 (m, 1H, O-CH_{2a}), 3.97 (dd, 1H, H-4, $J_{4,5}$ = 1.0 Hz), 3.86 (dd, 1H, H-2), 3.70 (s, 3H, $COOCH_3$), 3.68 - 3.62 (m, 1H, O-CH_{2b}), 3.45 (dd, 1H, H-5), 2.85 (dd, 1H, H-3e’, $J_{3e',3a'}$ = 13.2 Hz, $J_{3e',4'}$ = 5.1 Hz), 2.37, 2.29 (2s, 6H, 2 x NCO-CH_3), 2.16, 2.13, 2.03, 1.95 (4s, 12H, 4 x CO-CH_3), 1.95 (dd, 1H, H-3a’, $J_{3a',4'}$ = 10.6), 1.12 - 0.99 (m, 2H, O-$CH_2C\underline{H}_2$), 0.02 [s, 9H, $Si(CH_3)_3$]; ^{13}C NMR ($CDCl_3$): δ 174.4, 173.5 (2 x NCO), 170.6, 170.2, 170.1, 169.7 (4 x $\underline{C}O$-CH_3), 168.2 (C-1’), 138.0 (Cq, $\underline{C}_6H_5CH$), 128.9, 128.0, 126.8, 126.7, 126.5 (5 x $\underline{C}_6H_5CH$), 102.3 (C-1), 101.0 ($C_6H_5\underline{C}H$), 97.5 (C-2’), 75.4 (C-3), 74.1 (C-4), 70.0 (C-6’), 69.2 (C-6), 68.6 (C-2), 68.4 (C-8’), 66.9 (C-7’), 66.8 (C-4’), 66.6 (O-CH2), 66.2 (C-5), 62.0 (C-9’), 56.9 (C-5’), 53.0 ($COO\underline{C}H_3$), 39.1 (C-3’), 27.9, 26.0 (2 x NCO-$\underline{C}H_3$), 21.2, 20.9, 20.7 (x2) (4 x CO-$\underline{C}H_3$), 18.1 (O-$CH_2\underline{C}H_2$), -1.4 [$Si(CH_3)_3$].

1,2,4,6-Tetra-*O*-acetyl-*O*-{[methyl 5-(*N*-acetylacetamido)-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-*glycero*-α–D-*galacto*-non-2-ulopyranosid]-onate}-(2-3)-β– D-galactopyranosyl (2.25)

A solution of disaccharide **2.23** (2.35 g, 2.66 mmol) in acetic anhydride (10 mL) and pyridine (15 mL) was stirred at room temperature for 18 h and then quenched with methanol (10 mL), concentrated *in vacuo* and co-evaporated with toluene (3 x 5 mL) to afford **2.24,** as a white foam in a quantitative yield and used without purification (2.46 g). R_f 0.44 (acetone/toluene, 3/7, v/v); MALDI TOF *m/z* = 949 $[M + Na]^+$. To a solution of compound **2.24** in dry DCM (55 mL) under argon was added acetic anhydride (1.9 mL, 18.62 mmol). The solution was cooled to –5°C, and boron trifluoroetherate (843 µL, 6.65 mmol) was added dropwise. After 4 h at room temperature, the reaction mixture was diluted with DCM (30 mL) and washed with saturated aqueous $NaHCO_3$ (3x15 mL)and water(3x15 mL). The combined organic layers were dried ($MgSO_4$), filtered and the filtrate reduced under pressure. Purification of the crude material by silica gel column chromatography (eluent gradient: toluene to toluene/acetone, 3/7, v/v) gave compound **2.25** (α/β = 1/15) as a white foam in 89% yield (2.05 g). **2.25 (β):** R_f 0.33 (acetone/toluene, 3/7, v/v); MALDI TOF *m/z* = 886 $[M + Na]^+$; ^{1}H NMR($CDCl_3$) 600 MHz: δ 5.77 (d, 1H, H-1, $J_{1,2}$ = 8.5 Hz), 5.49 (ddd, 1H, H-4', $J_{4',5'}$ = 10.7 Hz), 5.45 (ddd, 1H, H-8', $J_{7',8''}$ = 9.8 Hz, $J_{8',9a'}$ = 3.1 Hz, $J_{8',9b'}$ = 6.2 Hz), 5.10 – 5.05 (m, 2H, H-7', H-2), 4.97 (dd, 1H, H-4, $J_{4,5}$ = 1.0 Hz), 4.70 (dd, 1H, H-3, $J_{2,3}$ = 10.2 Hz, $J_{3,4}$= 3.1 Hz), 4.55 (dd, 1H, H-6', $J_{5',6'}$ = 10.2 Hz, $J_{6',7'}$ = 1.8 Hz), 4.29 (dd, 1H, H-9a'), 4.27 (ρt, 1H, H-5'), 4.04 (dd, 1H, H-6a, $J_{5,6a}$ = 4.0 Hz, $J_{6a,6b}$ = 10.5 Hz), 3.99 (dd, 1H, H-6b, $J_{5,6b}$ = 7.1 Hz), 3.97 (ρt, 1H, H-5), 3.91 (dd, 1H, H-9b', $J_{9a',9b'}$ = 12.5 Hz, $J_{8',9b'}$ = 6.2 Hz), 3.84 (s, 3H, $COOCH_3$), 2.60 (dd, 1H, H-3e', $J_{3e',3a'}$ = 12.9 Hz, $J_{3e',4'}$ = 5.3 Hz), 2.29, 2.23 (2s, 6H, 2 x NCO-CH_3), 2.15, 2.11, 2.05, 2.03, 1.98, 1.89 (6s, 24H, 8 x CO-CH_3), 1.54 (dd, 1H, H-3a', $J_{3a',4'}$ = 10.3 Hz); ^{13}C NMR ($CDCl_3$): δ 174.2, 173.8 (2 x N-CO), 170.9, 170.7, 170.6, 170.4, 170.1, 170.0, 169.9, 169.1 (8 x $\underline{C}$O-CH_3), 168.0 (C-1'), 97.0 (C-2'), 92.2 (C-1), 71.7, 71.6 (C-5, C-3), 69.8 (C-6'), 69.2 (C-2), 68.0, 67.8, (C-8', C-4), 67.4, 67.2 (C-7', C-

4'), 62.7 (C-9'), 62.0 (C-6), 59.1 (C-5'), 53.4 (COO$\underline{C}H_3$), 38.6 (C-3'), 28.3, 27.0 (2 x NCO-$\underline{C}H_3$), 21.7, 21.2, 21.1, 20.9, 20.8 (8 x CO-$\underline{C}H_3$).

Methyl 2,4,6-tri-*O*-acetyl-*O*-{[methyl 5-(*N*-acetylacetamido)-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-*glycero*-α–D-*galacto*-non-2-ulopyranosid]-onate}-(2-3)-1-thio-β–D-galactopyranoside (2.26)

To a solution of **2.25** (2.05 g, 2.37 mmol) in dry DCM (80 mL), TMSSMe (2.6 mL, 5.93 mmol) and TMSOTf (1.0 mL, 3.56 mmol) was added and the reaction mixture left to stir at room temperature under argon. After 60 h the reaction mixture was diluted with DCM (30 mL). The solution was then washed with saturated aqueous $NaHCO_3$ (3 x 20 mL) and water (3 x 20 mL) and the combined organic layers dried ($MgSO_4$), filtered and concentrated *in vacuo*. Purification by silica gel column chromatography, using acetone/toluene (1/9 v/v) as the eluent, afforded **2.26** as a white foam in 92% yield (1.86 mg). R_f 0.38 (acetone/toluene, 3/7, v/v); MALDI TOF *m/z* = 874 $[M + Na]^+$; ^{1}H NMR ($CDCl_3$) 600 MHz: δ 5.54 (ddd, 1H, H-4', $J_{4',5'}$ = 10.9 Hz), 5.49 (m, 1H, H-8'), 5.17 (dd, 1H, H-7', $J_{6',7'}$ = 2.0 Hz, $J_{7',8'}$ = 9.8 Hz), 5.08 (ρt, 1H, H-2), 5.03 (d, 1H, H-4, $J_{4,5}$ = <1.0 Hz), 4.67 (dd, 1H, H-3, $J_{2,3}$ = 9.8 Hz, $J_{3,4}$ = 2.9 Hz), 4.61 (dd, 1H, H-6', $J_{5',6'}$ = 10.2 Hz), 4.55 (d, 1H, H-1, $J_{1,2}$ = 9.8 Hz), 4.32 (t, 1H, H-5'), 4.30 (dd, 1H, H-9a', $J_{8',9a'}$ = 3.7 Hz, $J9_{9a',9b'}$ = 12.6 Hz), 4.06 – 4.04 (m, 2H, H-6a, H-6b), 4.02 (dd, 1H, H-9b', $J_{8',9b'}$ = 4.9 Hz), 3.92 (ρt, 1H, H-5, $J_{5,6a}$ = $J_{5,6b}$ = 6.4 Hz), 3.89 (s, 3H, $COOCH_3$), 2.65 (dd, 1H, H-3e', $J_{3e',3a'}$ = 12.6 Hz, $J_{3e',4'}$ = 5.3 Hz), 2.35, 2.29 (2s, 6H, 2 x NCO-CH_3), 2.19 (s, 3H, SCH_3), 2.22, 2.17, 2.08, 2.04, 2.03, 1.95 (6s, 21H, 7 x CO-CH_3), 1.60 (dd, 1H, H-3a', $J_{3a',4'}$ = 10.0 Hz); ^{13}C NMR ($CDCl_3$): δ 174.3, 173.8 (2 x N-CO), 170.8, 170.7, 170.6, 170.4, 170.2, 170.1, 170.0 (7 x <u>C</u>O-CH_3), 168.1 (C-1'), 96.9 (C-2'), 83.1 (C-1), 74.5 (C-5), 72.6 (C-3), 69.7 (C-6'), 68.3, 68.0, 67.9, (C-4, C-2, C-8'), 67.3, 67.2 (C-7', C-4'), 62.5, 62.4 (C-9', C-6), 56.2 (C-5'), 53.3 (COO<u>C</u>H_3), 38.7 (C-3'), 28.3, 26.9 (2 x NCO-<u>C</u>H_3), 21.7, 21.3, 21.1, 21.0, 20.9 (7 x CO-<u>C</u>H_3), 11.6 (SCH_3).

3-(*N*-Benzoyloxycarbonyl) aminopropyl 3-*O*-(*p*-methoxybenzyl)-*O*-{[methyl 5-(*N*-acetylacetamido)-4,7,8,9-tetra-*O*-acetyl-3,5-D-dideoxy- *glycero*-α– D-*galacto*-non-2-ulopyranosid]-onate}-(2-3)-(2,4,6-tri-*O*-acetyl-β–D-galactopyranosyl)-(1-4)-6-*O*-benzyl-2-deoxy-2-acetamido-D-glucopyranoside (2.27)

A mixture of disaccharide **2.26** (284 mg, 0.333 mmol) and monosaccharide **2.22** (173 mg, 0.278 mmol) with activated crushed molecular sieves (500 mg, 4 Å) in dryDCM (10mL) was stirred for 2.5 h under argon at room temperature. NIS (133 mg, 0.590 mmol) and TMSOTf (5.0 μL, 0.028 mmol) were added and the reaction mixture stirred for 16 h at room temperature. The reaction mixture was diluted with DCM (20 mL), filtered and washed with aqueous $Na_2S_2O_3$ (15 mL,15%, w/v), saturated $NaHCO_3$ (2x10 mL)and H2O (2 x 10 mL). The organic phase was dried ($MgSO_4$) and filtered, and the filtrate was concentrated *in vacuo*. Purification by LH-20 size-exclusion column chromatography (DCM/methanol, 1/1, v/v) afforded **2.27** as a white foam (313 mg, 79%). R_f 0.30 (acetone/toluene, 2/3, v/v); MALDI TOF ***m/z*** **=** 1449 $[M + Na]^+$; ^{1}H NMR ($CDCl_3$) 500 MHz: δ 7.36 - 7.25 (m, 10H, $C_6\underline{H}_5CH_2$, $C_6\underline{H}_5CH_2OCO$), 7.18 (d, 2H, $OCH_3C_6H_4CH_2$, $J_{A,B}$ = 8.3 Hz), 6.84 (d, 2H, $OCH_3C_6H_4CH_2$), 6.09 (d, 1H, N<u>H</u>-CO-CH_3, $J_{NH,2}$ = 8.3 Hz), 5.42 - 5.38 (m, 1H, NH-COO), 5.59 – 5.53 (m, 2H, H-4'', H-8''), 5.16 (dd, 1H, H-7'' , $J_{6'',7''}$ = 2.4 Hz, $J_{7'',8''}$ = 8.8 Hz), 5.07 (s, 2H, $C_6H_5C\underline{H}_2OCO$), 5.00 – 4.97 (m, 2H, H-2', H-4’), 4.78 (d, 1H, H-1’, $J_{1',2'}$ = 7.8 Hz), 4.69 (dd, 1H, H-3’, $J_{2',3'}$ = 10.1 Hz, $J_{3',4'}$ = 2.2 Hz), 4.65 (d, 1H, $C_6H_5C\underline{H}_{2a}$, $J_{CH2a,CH2b}$ = 11.9 Hz), 4.62 (dd, 1H, H-6’’), 4.60 (d, 1H, H-1, $J_{1,2}$ = 8.6 Hz), 4.58 (d, 1H, $OCH_3C_6H_4C\underline{H}_{2a}$, $J_{CH2a,CH2b=}$ 12.5 Hz), 4.55 (d, 1H, $C_6H_5C\underline{H}_{2b}$), 4.51 (d, 1H, $OCH_3C_6H_4C\underline{H}_{2b}$), 4.34 (t, 1H, H-5’’), 4.29 (d, 1H, H-9a’’ , $J_{9a'',9b''}$ = 12.1 Hz, $J_{8'',9a''}$ = <1.0 Hz), 4.05 (d, 1H, H-6a’, $J_{5',6a'}$ = <1.0 Hz, $J_{6a',6b'}$ = 9.0 Hz), 3.98 - 3.93 (m, 3H, H-3, H-6b’, H-9b’’), 3.90 (s, 3H, $COOCH_3$), 3.85 – 3.71 (m, 7H, H-2, H-5’, O-CH_{2a}, H-6a, H-6b, H-5, H-4), 3.79 (s, 3H, OCH_3), 3.51 - 3.46 (m, 1H, OCH_{2b}), 3.31 - 3.29 (m, 1H, CH_{2a}-N), 3.24 – 3.21 (m, 1H, CH_{2b}-N), 2.66 (dd, 1H, H-3e’’, $J_{3e'',3a''}$ = 12.7 Hz, $J_{3e''4''}$ = 5.4 Hz), 2.37, 2.31 (2s, 6H, 2 x NCO-CH_3), 2.20, 2.16, 2.07, 2.04, 1.96 (x 2), 1.95, 1.93 (7s, 24H, 7 x CO-CH_3, NH-CO-$C\underline{H}_3$), 1.78 - 1.70 (m, 2H, O-$CH_2C\underline{H}_2CH_2$), 1.62 (t, 1H, H-3a’’); ^{13}C NMR ($CDCl_3$): δ 174.3, 173.9 (2 x N-$\underline{C}$O-CH_3), 170.9 – 169.4 (7 x $\underline{C}$O-CH_3, NH-$\underline{C}$O-CH_3), 168.2 (C-1’’), 159.4 (2 x Cq, $OCH_3\underline{C}_6H_4CH_2$), 156.7 (NH-COO), 138.3 (Cq, $\underline{C}_6H_5CH_2$), 137.0 (Cq, $\underline{C}_6H_5CH_2$ OCO), 130.5 (2 x $OCH_3\underline{C}_6H_4CH_2$), 128.6 - 127.5 (5 x $\underline{C}_6H_5CH_2$, 5 x $\underline{C}_6H_5CH_2OCO$), 113.9 (2 x $OCH_3\underline{C}_6H_4CH_2$), 100.6 (C-1), 99.8 (C-1’), 97.1 (C-2’’), 76.6 (C-4), 74.5 (C-5), 73.1 ($OCH_3C_6H_4\underline{C}H_2$), 72.4 ($C_6H_5\underline{C}H_2$), 71.2 (C-3’), 70.9 (C-2’), 70.8 (C-5’), 69.7 (C-3), 69.6 (C-6’’), 69.5 (C-6), 67.5 (C-4’), 67.3 (x 2) (C-8’’, C-4’’), 66.7 (C-7’’), 66.5 ($C_6H_5\underline{C}H_2OCO$), 67.4 (O-$CH_2$), 62.5 (C-9’’), 61.8 (C-6’), 56.4 (C-2), 56.2 (C-5’’), 55.4 (OCH_3), 53.2 ($COO\underline{C}H_3$), 38.7 (CH_2-N), 37.0 (C-3’’), 29.5 (O-$CH_2\underline{C}H_2CH_2$), 28.1, 26.8 (2 x NCO-$\underline{C}H_3$), 23.3, 21.4, 21.0, 20.9, 20.8, 20.7 (x 2) (7 x CO-$\underline{C}H_3$, NHCO-$\underline{C}H_3$).

3-(*N*-Benzoyloxycarbonyl) aminopropyl *O*-{[methyl 5-(*N*-acetylacetamido)-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-*glycero*-α–D-*galacto*-non-2-ulopyranosid]-onate}-(2-3)-(2,4,6-tri-*O*-acetyl-β–D-galactopyranosyl)-(1-4)-6-*O*-benzyl-2-acetamido-2-deoxy-β–D-glucopyranoside (2.28)

To a stirred solution of compound **2.27** (313 mg, 0.219 mmol) in dry DCM (9 mL) was added TFA (1 mL). After stirring at 6 h at room temperature the reaction mixture was diluted with DCM (20 mL), poured into ice-water (50 mL) and the organic layer washed with saturated aqueous $NaHCO_3$ (2x10 mL)and water (2x10 mL). The combined filtrates were dried ($MgSO_4$), filtered, concentrated *in vacuo* and purified by silica gel column chromatography (eluent gradient: toluene to toluene/acetone, 1/1, v/v) to give **2.28** as a white foam (212 mg, 74%). R_f 0.16 (acetone/toluene, 2/3, v/v); MALDI TOF *m/z* = 1329 $[M + Na]^+$; ^1H NMR ($CDCl_3$) 500 MHz: δ 7.34 - 7.26 (m, 10H, $C_6\underline{H}_5CH_2$, $C_6\underline{H}_5CH_2OCO$), 6.23 (d, 1H, N$\underline{H}$-$COCH_3$, $J_{NH,2}$= 7.8 Hz), 5.54 (ddd, 1H, H-4'' , $J_{3a'',4''}$ = 5.4 Hz, $J_{4'',5''}$ = 10.7 Hz), 5.50 (ddd, 1H, H-8'', $J_{8'',9b''}$ = 6.4 Hz), 5.20 - 5.18 (m, 1H, NH-COO), 5.12 (dd, 1H, H-7'', $J_{6'',7''}$ = 2.4 Hz, $J_{7'',8''}$ = 8.8 Hz), 5.08 (s, 2H, $C_6H_5C\underline{H}_2OCO$), 5.01 (t, 1H, H-2'), 4.93 (d, 1H, H-4', $J_{4',5'}$ = <1.0 Hz), 4.77 (d, 1H, H-1, $J_{1,2}$ = 7.8 Hz), 4.67 (dd, 1H, H-3', $J_{2',3'}$ = 10.3 Hz, $J_{3',4'}$ = 3.4 Hz), 4.64 – 4.55 (d, 3H, H-6'' , $C_6H_5C\underline{H}_{2a}$, $J_{CH2a,CH2b}$= 11.9 Hz), 4.44 (d, 1H, H-1', $J_{1',2'}$ = 8.3 Hz), 4.32 (ρt, 1H, H-5''), 4.26 (d, 1H, H-9a'' , $J_{9a'',9b''}$ = 11.7 Hz, $J_{8'',9a''}$ = <1.0 Hz), 4.02 (dd, 1H, H-6a', $J_{5',6a'}$ = 3.4 Hz, $J_{6a',6b'}$ = 10.3Hz), 4.05 – 3.98 (m, 2H, H-5', H-6b'), 3.93 - 3.89 (m, 2H, H-9'', OCH_{2a}), 3.89 (s, 3H, $COOCH_3$), 3.78 - 3.73 (m, 2H, H-6a, H-3), 3.69 - 3.62 (m, 2H, H-2, H-6b), 3.53 – 3.45 (m, 3H, H-4, H-5, OCH_{2b}), 3.16 - 3.10 (m, 2H, CH_2-N), 2.65 (dd, 1H, H-3e'', $J_{3e'',3a''}$ = 12.7 Hz), 2.36, 2.30 (2s, 6H, 2 x NCO-CH_3), 2.16, 2.15, 2.09, 2.07, 2.02, 1.20, 1.94, 1.95 (8s, 24H, 7 x CO-CH_3, NH-CO-C$\underline{H}_3$), 1.82 - 1.76 (m, 1H, O-$CH_2C\underline{H}_{2a}CH_2$), 1.65 - 1.57 (m, 2H, O-$CH_2C\underline{H}_{2b}CH_2$, H-3a''); ^{13}C NMR ($CDCl_3$): δ 174.1, 173.8 (2 x N-$\underline{C}$O-CH_3), 171.5 (NH-$\underline{C}$O-CH_3), 170.8, 170.5, 170.3, 170.1, 170.0, 169.8 (7 x $\underline{C}$O-CH_3), 168.1 (C-1''), 157.0 (NH-COO), 138.7 (Cq, $\underline{C}_6H_5CH_2$), 136.9 (Cq, $\underline{C}_6H_5CH_2OCO$), 128.5 - 127.5 (5 x $\underline{C}_6H_5CH_2$, 5 x $\underline{C}_6H_5CH_2OCO$), 101.6 (C-1'), 101.1 (C-1), 96.9 (C-2''), 82.5 (C-4), 74.4 (C-5), 73.6 (x 2) (C-3, $C_6H_5\underline{C}H_2$), 71.7 (C-3'), 71.1 (C-5'), 70.3 (C-2'), 69.8 (C-6''), 69.4 (C-6), 67.9 (C-8''), 67.7 (C-4'), 67.4 (C-7''), 67.1 (C-4''), 66.8 ($C_6H_5\underline{C}H_2OCO$), 66.6 (O-$CH_2$), 62.5 (C-9''), 62.3 (C-6'), 56.4 (C-2), 56.0 (C-5''), 53.3 (OCH_3), 38.7 (CH_2-N), 38.6 (C-3''), 29.9 (O-$CH_2\underline{C}H_2CH_2$), 28.2, 26.8 (2 x NCO-$\underline{C}H_3$), 23.6 (NH-CO-$\underline{C}H_3$), 21.5, 21.1, 20.8 (x 2), 20.7, 20.5 (x 2) (7 x CO-$\underline{C}H_3$).

3-aminopropyl *O*-[sodium (5-acetamido-3,5-dideoxy-D-*glycero*-α-D-*galacto*-non-2-ulopyranosid) onate]-(2-3)-*O*-β-D-galactopyranosyl-(1-4)-2-acetamido-2-deoxy-β-D-glucopyranoside (2.29deprotected)

Compound **2.28** (89 mg, 68 μmol) was dissolved in methanol (1.5 mL) and 1M aqueous NaOH (0.5 mL) added. After stirring for 40 h the mixture was purified by Sephadex G-15 size-exclusion column chromatography to give **2.29** as a white solid (66 mg, 99 %). Residue **2.29** (46mg, 47 μmol) was dissolved in 50% aqueous ethanol (2 mL) and acetic acid (0.4 mL), and Pd/C (20 mg) was added. The reaction mixture was stirred under hydrogen for 4 h. The catalyst was then filtered off and the filtrate concentrated under reduced pressure. The residue was dissolved in water (1 mL) and 1M aqueous NaOH (0.2 mL). After stirring for 1 h, the reaction mixture was purified by Sephadex G-15 size-exclusion column chromatography, to afford deprotected **2.29d** as a white powder

(30 mg, 85 %). MALDI TOF *m/z* = 777 $[M + Na]^+$; ^{1}H NMR (D_2O) 500 MHz: δ 4.58 (d, 1H, H-1’, $J_{1',2'}$ = 7.8 Hz), 4.54 (d, 1H, H-1, $J_{1,2}$ = 8.3 Hz), 4.14 (dd, 1H, H-3’, $J_{2',3'}$ = 9.8 Hz, $J_{3',4'}$ = 2.9 Hz), 4.07 - 4.03 (m. 2H, H-6a, OCH_{2a}), 3.98 (d, 1H, H-4’, $J_{4',5'}$ = <1.0 Hz), 3.93 – 3.83 (m, 4H, H-5’’, H-8’’, H-9a’’, H-6b), 3.80 - 3.70 (m, 8H, H-2, H-3, H-4, H-5’, H-6a’, H-6b’, H-4’’, OCH_{2b}), 3.70 – 3.61 (m, 4H, H-5, H-6’’, H-7’’, H-9b’’), 3.60 (dd, 1H, H-2’), 3.10 (t, 2H, CH_2-N), 2.78 (dd, 1H, H-3e’’, $J_{3e''3a''}$ = 12.7 Hz, $J_{3e'',4''}$ = 4.9 Hz), 2.07, 2.06 (2s, 6H, 2 x NH-CO- CH_3), 1.98 - 1.95 (m, 2H, O-$CH_2C\underline{H_2}CH_2$), 1.82 (t, 1H, H-3a’’); ^{13}C NMR (D2O): δ 175.2, 174.9 (2 x CO-CH_3), 174.0 (C-1’’), 102.6 (C-1’), 101.1 (C-1), 99.0 (C-2’’), 78.3 (C-4), 75.3 (C-3’), 75.0 (C-5’), 74.5 (C-5), 72.9 (C-6’’), 72.3 (C-3), 71.6 (C-8’’), 69.2 (C-2’), 67.9 (C-7’’), 67.8, (C-4’’), 67.7 (O-CH_2), 67.2 (C-4’), 62.6 (C-9’’), 60.9 (C-6’), 59.8 (C-6), 54.9 (C-2), 51.5 (C-5’’), 39.5 (C-3’’), 37.3 (CH_2-N), 26.4 (O-$CH_2\underline{C}H_2CH_2$), 21.8 (2 x NHCO-$\underline{C}H_3$).

3-acetamidopropyl *O*-[sodium (5-acetamido-3,5-dideoxy-D-*glycero*-α–D-*galacto*-non-2-ulopyranosid) onate]-(2-3)-*O*-(β–D-galactopyranosyl)-(1-4)-2-acetamido-2-deoxy-β–D-glucopyranoside (2.1)

Compound **2.29d** (10.5 mg, 13.9 μmol) was dissolved in methanol (1 mL) followed by the addition of acetic anhydride (0.5 mL). After stirring for 16 h at room temperature, the reaction mixture was concentrated *in vacuo* and the residue co-evaporated with toluene (3 x 5 mL). The residue was dissolved in water (1 mL) and 1M aqueous NaOH (0.2 mL). After stirring for 1 h, the reaction mixture was purified by Sephadex G-15 size-exclusion column chromatography to furnish monovalent **2.1** as a glassy solid (10.4 mg, 94%). MALDI TOF *m/z* = 819 $[M + Na]^+$; ^{1}H NMR (D_2O) 500 MHz: δ 4.58 (d, 1H, H-1’, $J_{1',2'}$

= 7.8 Hz), 4.54 (d, 1H, H-1, $J_{1,2}$ = 8.3 Hz), 4.15 (dd, 1H, H-3', $J_{2',3'}$ = 9.8 Hz, $J_{3',4'}$ = 2.9 Hz), 4.03 (dd, 1H, H-6a, $J_{5,6a}$ = <1.0 Hz, $J_{6a,6b}$ = 10.7 Hz), 3.99 (d, 1H, H-4', $J_{4',5'}$ = <1.0 Hz), 3.97 – 3.86 (m, 5H, OCH_{2a}, H-5'', H-8'', H-9a'', H-6b), 3.80 - 3.70 (m, 7H, H-2, H-3, H-4, H-5', H-6a', H-6b', H-4''), 3.70 – 3.61 (m, 5H, OCH_{2b}, H-5, H-6'', H-7'', H-9b''), 3.60 (dd, 1H, H-2'), 3.31 – 3.25 (m, 1H, CH_{2a}-N), 3.22 – 3.16 (m, 1H, CH_{2b}-N), 2.79 (dd, 1H, H-3e'', $J_{3e'',3a''}$ = 12.7 Hz, $J_{3e'',4''}$ = 4.9 Hz), 2.06 (x 2) (s, 6H, 2 x NH-CO-C$\underline{H}_3$), 2.01 (s, 3H, NH-CO-C$\underline{H}_3$), 1.83 (t, 1H, H-3a''), 1.82 - 1.77 (m, 2H, O-$CH_2C\underline{H}_2CH_2$); ^{13}C NMR (D_2O): δ 175.2, 174.4, 174.2 (3 x $\underline{C}$O-CH_3), 174.0 (C-1''), 102.4 (C-1'), 101.0 (C-1), 99.7 (C-2''), 78.5 (C-4), 75.6 (C-3'), 75.3 (C-5'), 74.9 (C-5), 73.0 (C-6''), 72.5 (C-3), 71.8 (C-8''), 69.5 (C-2'), 68.3 (C-4''), 68.1 (C-7''), 67.9 (O-CH_2), 67.7 (C-4'), 62.7 (C-9''), 61.1 (C-6), 60.3 (C-6'), 55.1 (C-2), 51.9 (C-5''), 39.6 (C-3''), 36.3 (CH_2-N), 28.3 (O-$CH_2\underline{C}H_2CH_2$), 22.1 (x 2), 21.9 (3 x NH-CO-$\underline{C}H_3$).

Polymeric substrate (2.2)

As previously described, to a solution of **2.29d** (5.1mg, 6.7μmol) and poly[*N*-*(*acryloxy)succinimide] (6.0mg, 35.5 μmol of *N*-hydroxysuccinimide ether, 19% theoretical loading) in DMF (500 μL) was added DIPEA (3.5 μl, 20.3 μmol) (193). The solution was stirred at room temperature for 18 h, heated at 65^{o}C for 5 h, and then cooled to room temperature. An aqueous solution of NH_4OH (20%, 2.0 mL) was added and the mixture stirred at room temperature for a further 18 h. The resulting mixture was dialyzed against distilled water for 3 days and then lyophilized to afford **2.2** as a fluffy white foam (6.9 mg, 95%). Selected ^{1}H NMR data (D_2O) 500 mHZ: δ 4.56 (dd, 2H, H-1', H-1), 4.14

(d, 1H, H-3’), 3.98 (s, 1H, H-4’), 3.31 – 3.16 (m, 2H, CH_2-N), 2.79 (dd, 1H, H-3e’’, $J_{3e'',3a''}$ = 12.7 Hz, $J_{3e'',4''}$ = 4.9 Hz), 2.06 (s, 6H, 2 x NHCO-C<u>H</u>$_3$).

S-[Methyl (5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-*glycero*-α-D-*galacto*-non-2-ulopyranosyl)onate]-(2→3)-1,2:5,6-di-*O*-isopropylidene-3-thio-α-D-galactofuranose 2.33

Diethylamine (6.25 mL) was added to a stirring solution of triflate **2.31** (194) (1.14 g, 2.9 mmol) and thioacetate **2.32** (39) (1.22 g, 2.22 mmol) in DMF (12.5 mL) at 0°C. The reaction mixture warmed to room temperature over a 6 hour period. The mixture was concentrated, dissolved in EtOAc (125 mL) washed with 1M HCl, saturated $NaHCO_3$ (3 x 50 mL) and water (3 x 50 mL). The organic filtrate was concentrated *in vacuo* and purfied by flash chromatography (eluent: hexane–EtOAc, 7 : 3) gave compound **2.33** as an amorphous solid (1.13 g, 67%). The product was used in subsequent reactions without further purification. MALDI TOF *m/z* = 772 $[M + Na]^+$; ^{1}H NMR ($CDCl_3$) 500 MHz: δ 1.30, 1.34, 1.40, 1.56 (4s, 4 × 3H, 2 × $C(CH_3)_2$), 1.87 (m, 4 H, AcN, H-$3b_a$), 1.99, 2.00, 2.05, 2.10 (4 × 3 H, 4 s, 4 × AcO), 2.76 (dd, 1 H, $J_{3ba,3be}$ = 12.5 Hz , $J_{3be,4b}$ = 4.7 Hz, H-$3b_e$), 3.56 (m, 1 H, H-3_a), 3.59 (dd, 1 H, $J_{5a,6a'}$ = 6.6 Hz, $J_{6a,6a'}$ = 8.5 Hz, H-$6_{a'}$), 3.74–3.83 (m, 5 H, H-4a, H-6b, CO_2CH_3), 3.94 (q, 1 H, $J_{4b,5b}$ = $J_{5b,6b}$ = $J_{5b,NH}$ = 10.2 Hz, H-5b), 4.06–4.14 (m, 2 H, 6a-, H-9b), 4.23 (dd, 1 H, $J_{8b,9b'}$ = 2.5 Hz, $J_{9b,9b'}$ = 12.6 Hz, H-9b’), 4.43 (m, 1 H, H-5a), 4.76 (dd, 1 H, $J_{1a,2a}$ = 3.6 Hz, $J_{2a,3a}$ = 0.8 Hz, H-2a), 4.91 (m, 1 H, 4b-H), 5.29 (1 H, dd, J6b,7b 1.6, J7b,8b 9.3, 7b-H), 5.34 (1 H, d, J5b,NH, NH), 5.58 (1 H, m, H-8b), 5.82 (d, 1 H, $J_{1a,2a}$, H-1a-H).

S-[Methyl (5-acetamido-4,7,8,9-tetra-O-acetyl-3,5-dideoxy-D-*glycero*-α-D-*galacto*-non-2-ulopyranosyl)onate]-(2→3)-3-thio-D-galactose (2.34)

Compound **2.33** (200 mg, 266 µmol) was dissolved in aqueous TFA (2 mL) and the solution was stirred at room temperature for 15 min followed by concentration. Flash chromatography (eluent; EtOAc then MeOH) gave compound **2.34** (125 mg, 69%); MALDI TOF *m/z* = 692 $[M + Na]^+$; ^{1}H NMR (pyridine-d5) 500 MHz : δ 5.64 (d, $J_{1,2}$ = 4.4 Hz, H-1a α-pyr/fur), 5.69 (d, J1,2 7.4, 1a-H β-pyr), 5.88 (d, $J_{1,2}$ = 4.4 Hz, H-1a α-fur/pyr), 5.90 (s, H-1a β-fur).

S-[Methyl (5-acetamido-4,7,8,9-tetra-O-acetyl-3,5-dideoxy-D-*glycero*-α-D-*galacto*-non-2-ulopyranosyl)onate]-(2→3)-1,2,4,6-tetra-O-acetyl-3-thio-α/β-D-galactopyranose and -(2→3)-1,2,5,6-tetra-O-acetyl-3-thio-α/β-D-galactofuranose (2.35)

2.34 (50 mg, 74 µmol) was dissolved in pyridine (0.5 mL) and acetic anhydride (0.5 mL) and stirred at room temperature overnight. The mixture was concentrated *in vacuo* and purified by flash chromatography (eluent gradient: DCM–acetone, 9 : 1→3 : 1) gave the **2.34** (53 mg, 85%) as an α/β mixture in both pyranose and furanose ring forms. MALDI TOF *m/z* = 860 $[M + Na]^+$; ^{1}H NMR ($CDCl_3$) 500 MHz: δ 2.60–2.76 (m, 1 H, H-$3b_e$), 3.83 (s, 3 H, CO_2CH_3), 5.64 (ddd, $J_{7b,8b}$ = 9.6 Hz, $J_{8b,9b}$ = 2.5 Hz, $J_{8b,9b'}$ = 6.9 Hz, H-8b β-pyr), 6.03 (d, $J_{1,2}$ = 8.2 Hz, H-1a β-pyr), 6.17 (s, H-1a β-fur), 6.23 (d, $J_{1,2}$ = 4.1 Hz, H-1a α-pyr), 6.25 (d, $J_{1,2}$ = 4.1 Hz, H-1a α-fur); ^{13}C NMR ($CDCl_3$): δ 36.8 (C-3b β-pyr), 37.3 (C-3b α-pyr), 37.4 (C-3b β-fur), 38.0 (C-3b α-fur), 80.45 (C-2b α-pyr), 80.8 (C-2b β-pyr), 82.6 (C-2b α-fur), 83.25 (C-2b β-fur), 89.1 (C-1a α-pyr), 92.7 (C-1a β-pyr), 92.8 (C-1a α-fur), 99.4 (C-1a β-fur) .

S-[Methyl (5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-*glycero*-α-D-*galacto*-non-2-ulopyranosyl)onate]-(2→3)-2,4,6-tri-*O*-acetyl-3-thio-β-D-galactopyranose (2.36)

Compound **2.35** (105 mg, 0.125 mmol) and hydrazinium acetate (12.5 mg, 0.13 mmol) in DMF (3.5 ml) was heated to 50^{o}C for 7 h. After concentrating *in vacuo,* the reaction mixture was dissolved in DCM and washed with 1M HCl, saturated $NaHCO_3$ (3 x 50 mL) and water (3 x 50 mL). Followed by concentrating, the product was purified via flash chromatography (eluent; toluene–MeOH, 9 : 1) to afford compound **2.35** as a syrup (70 mg, 69%). MALDI TOF *m/z* = 818 [M + Na]$^+$; ^{1}H NMR ($CDCl_3$) 500 MHz: δ 1.86 (s, 3 H, AcN), 1.88 (dd, 1 H, $J_{3a,3e}$ = 12.6 Hz, $J_{3a,4}$ = 11.8 Hz, H-3b$_a$), 2.01, 2.03, 2.04, 2.06, 2.07, 2.19, 2.22 (7 × 3 H, 7 s, 7 × AcO), 2.64 (dd, 1 H, $J_{3ba,3be}$ = 12.6 Hz, $J_{3be,4b}$ = 4.7 Hz, H-3b$_e$), 3.66–3.74 (m, 2 H, H-3a, H-6b), 3.83 (s, 3 H, CO_2CH_3), 3.93–4.19 (m, 5 H, H-5a, 6-H_2, H-5b, H-9b), 4.29 (dd, 1 H, $J_{8b,9b'}$ = 2.7 Hz, $J_{9b,9b'}$ = 12.4 Hz, H-9b’), 4.70 (dd, 1 H, $J_{1a,2a}$ = 7.7 Hz, $J_{2a,3a}$ = 11.8 Hz, H-2a), 4.83 (m, 1 H, H-4b), 4.91 (dd, 1 H, $J_{3a,4a}$ = 3.6 Hz, $J_{4a,5a}$ = 1.1 Hz, H-4a), 5.04 (d, 1 H, $J_{1a,2a}$, H-1a), 5.22 (d, 1 H, $J_{5b,NH}$ = 10.2 Hz, NH), 5.31 (dd, 1 H, $J_{6b,7b}$ = 2.2 Hz, $J_{7b,8b}$ = 10.2 Hz, H-7b), 5.59 (m, 1 H, H-8b).

S-[Methyl (5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-g*lycero*-α-D-galacto-non-2-ulopyranosyl)onate]-(2→3)-2,4,6-tri-*O*-acetyl-3-thio-α/β-D-galactopyranosyl trichloroacetimidate (2.37)

DBU (20 µl) was added to a stirred solution of **2.36** (32 mg, 34 µmol) in DCM (1.25 mL) and trichloroacetonitrile (0.25 ml). After 2 h at room temperature the solution was concentrated *in vacuo* and passed down a silica plug (eluent; DCM then acetone). MALDI TOF *m/z* = 962 $[M + Na]^+$; ^{1}H NMR ($CDCl_3$) 500 MHz: δ 1.85 (s, 3 H, AcN), 1.88 (m, 1 H, H-3b_a), 1.99, 2.00, 2.02, 2.05, 2.09, 2.14, 2.21 (7 s, 7 × 3 H, 7 × AcO), 2.64 (dd, 1 H, $J_{3ba,3be}$ = 12.6 Hz, $J_{3be,4b}$ = 4.4 Hz, H-3b_e), 3.70 (dd, 1 H, $J_{5b,6b}$ = 10.7 Hz, $J_{6b,7b}$ = 2.2 Hz, H-6b), 3.81 (dd, 1 H, $J_{2a,3a}$ = 11.5 Hz, $J_{3a,4a}$ = 3.6 Hz, H-3a), 3.84 (s, 3 H, CO_2CH_3), 3.93 (dd, 1 H, $J_{8b,9b}$ = 6.0 Hz, $J_{9b,9b'}$ = 12.4 Hz, H-9b), 4.01 (dd, 1 H, $J_{5a,6a}$ = 6.9 Hz, $J_{6a,6a'}$ =11.3 Hz, H-6a), 4.02–4.12 (m, 2 H, 6a'-, H-5b), 4.30 (dd, 1 H, $J_{8b,9b'}$ = 2.2 Hz, $J_{9b,9b'}$, H-9b'), 4.33 (m, 1 H, H-5a), 4.82 (m, 1 H, m, H-4b), 4.95 (d, 1 H, $J_{3a,4a}$, H-4a), 5.10 (dd, 1 H, $J_{1a,2a}$ = 8.0 Hz, $J_{2a,3a}$, H-2a), 5.16 (d, 1 H, $J_{5b,NH}$ = 10.2 Hz, NH), 5.28 (dd, 1 H, $J_{6b,7b}$, $J_{7b,8b}$ = 10.2 Hz, H-7b), 5.62 (ddd, 1 H, $J_{7b,8b}$, $J_{8b,9b}$ = 2.2 Hz, $J_{8b,9b'}$, H-8b), 6.15 (d, 1 H, $J_{1a,2a}$, H-1a), 8.43 (s, 1 H, C=NH).

3-(*N*-Benzoyloxycarbonyl) aminopropyl S-[methyl(5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-*glycero*-α-D-*galacto*-non-2-ulopyranosyl)onate]-(2→3)-2,4,6-tri-*O*-acetyl-3-thio-β-D-galactopyranose (2.38)

A suspension of **2.37** (25 mg, 53 µmol) with activated and crushed 4Å MS (100 mg) in dry DCM (1 mL) was stirred for 3 h at room temperature. Following cooling to 0^{o}C, aminopropyl spacer (65 µmol) was added then boron trifluoride-diethyl ether and was stirred for 3 h. The mixture was diluted in DCM (4 mL) and filtered through Celite. The filtrate was washed with 1M HCl, saturated $NaHCO_3$ (3 x 2 mL) and water (3 x 2 mL). After concentration *in vacuo*, the product was purified via silica gel column (eluent: DCM:MeOH, 97:3) giving **2.38** (16 mg, 60%). MALDI TOF *m/z* = 904 $[M + Na]^+$; ^1H NMR ($CDCl_3$) 500 MHz: δ 1.61 – 1.67 (m, 2 H, O-CH_2-C<u>H</u>$_2$-CH_2), 1.84 (m, 4 H, AcN, H-3b_a), 2.00 – 2.23 (6 s, 21 H, 7 × AcO), 2.60 (dd, 1 H, $J_{3ba,3be}$ = 12.4 Hz, $J_{3be,4b}$ = 4.9 Hz, H-3b_e), 3.17 – 3.09 (m, 2 H, CH_2-N), 3.66 (dd, 1 H, $J_{2a,3a}$ = 10.9 Hz, $J_{3a,4a}$ = 3.2 Hz, H-3a), 3.70 (dd, 1 H, $J_{5b,6b}$ = 10.6 Hz, $J_{6b,7b}$ = 2.2 Hz, H-6b), 3.78 – 3.88 (m, 3 H, CO_2CH_3), 3.96 – 4.08 (m, 5 H, H-5a, H_2-6a, H-5b, H-9b), 4.29 (dd, 1 H, $J_{8b,9b}$ = 2.8 Hz, $J_{9b,9b'}$ = 12.4 Hz, H-9b'), 4.73 – 4.88 (m, 3 H, H-1a, H-2a, H-4b), 4.88 (d, 1 H, $J_{3a,4a}$ = 3.0 Hz, H-4a), 5.16 (d, 1 H, $J_{5b,NH}$ = 10.0 Hz, NH), 5.33 (dd, 1 H, $J_{6b,7b}$, $J_{7b,8b}$ = 10.3 Hz, H-7b), 5.55 (ddd, 1 H, $J_{7b,8b}$, $J_{8b,9b'}$ = 2.5 Hz, $J_{8b,9b}$ = 5.4 Hz, H-8b), 6.22 (d, 1 H, N<u>H</u>-$COCH_3$, $J_{NH,2}$ = 7.7 Hz), 7.28 – 7.40 (m, 25 H, 4 × C_6<u>H</u>$_5$$CH_2$, C_6<u>H</u>$_5$$CH_2$OCO).

3-aminopropyl S-[(5-acetamido-3,5-dideoxy-D-*glycero*-α-D-*galacto*-non-2-ulopyranosyl)onate]-(2→3)-3-thio-β-D-galactopyranose (2.39)

Compound **2.38** (16 mg) was stirred in MeOH (0.5 mL) and 1M aqueous NaOH (0.1 mL) for 24 h. After stirring, the mixture was put directly onto a Sephadex G-15 SEC column to give **2.39** (6 mg, 59%). MALDI TOF *m/z* = 566 $[M + Na]^+$; ^1H NMR (DC_2O) 500 MHz: δ 1.74 (m, 1 H, H-3b_a), 1.94 – 1.99 (m, 2 H, O-$CH_2CH_2CH_2$), 2.04 (s, 3 H, AcN), 2.94 (dd, 1 H, $J_{3ba,3be}$ = 12.5 Hz, $J_{3be,4b}$ = 4.0 Hz, H-3b_e), 3.33 (dd, 1 H, $J_{1a,2a}$ = 7.6 Hz, $J_{2a,3a}$ = 11.7 Hz, H-2a), 3.59 (dd, 1 H, $J_{8b,9b}$ = 11.9 Hz, H-9b), 3.70 (dd, 1 H, $J_{5a,6a}$ = 5.5

Hz, $J_{6a,6a'}$ = 11.1 Hz, H-6a), 3.66 – 3.78 (m, 3 H, H-6a’, H-4b, H-5b), 4.31 (d, 1 H, $J_{1a,2a'}$,

H-1a).

Chapter 3

Selective Recognition of Synthetic Lysine and *meso* Diaminopimelic Acid-Type Peptidoglycan Fragments By Human Peptidoglycan Recognition Protein-Iα and –S [1]

[1] * This work was supported by National Institutes of Health Grants GMO65248 (to G.-J. B.) and AI47990 and AI065612 (to R. A. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

3.1 Abstract

The interactions of a range of synthetic peptidoglycan derivatives with PGRP-Iα and PGRP-S have been studied in real time using Surface Plasmon Resonance (SPR). A dissociation constant of K_D = 62 μM was obtained for the interaction of peptidoglycan recognition protein (PGRP)-Iα with the lysine containing muramyl pentapeptide (compound 6). The normalized data for the lysine-containing muramyl tetra- (compound 5) and pentapeptide (compound 6) showed that that these compounds have similar affinities whereas a much lower affinity for muramyl tripeptide (compound 3) was measured. Similar affinities were obtained when the lysine moiety of the muramyl peptides was replaced by *meso*-diaminopimelic acid (DAP). Furthermore, the compounds that contained only a stem peptide (pentapeptide, compound 1) and (DAP-PP, compound 2) as well as muramyl dipeptide (compound 3) exhibited no binding indicating that the muramyl tripeptide (compound 4) is the smallest peptidoglycan fragment that can be recognized by PGRP-Iα. Surprisingly, PGRP-S derived significantly higher affinities for the DAP-containing fragments to similar lysine containing derivatives and the following dissociation constants were measured: muramylpentapeptide-DAP K_D =104 nM, muramyltetrapeptide-DAP 92.4 nM, and muramyltripeptide-DAP 326 nM. The binding profiles were rationalized by using a recently reported X-ray crystal structure of PGRP-Iα with the lysine containing muramyl tripeptide (compound 4).

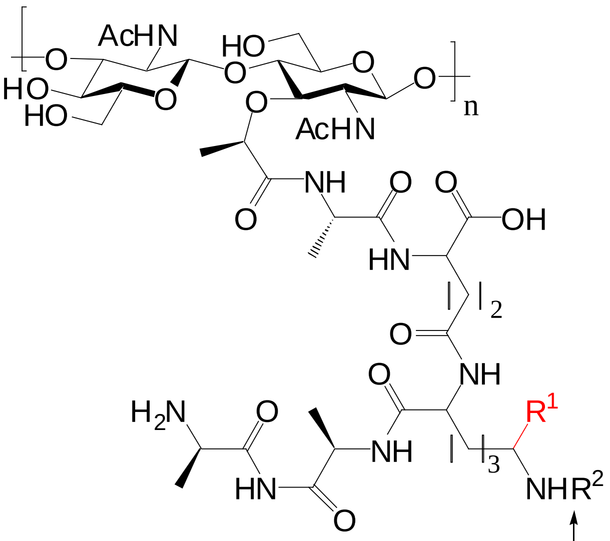
3.2 Introduction

The innate immune system is an ancient evolutionary system of defense against microbial infections (195-198). It responds rapidly to highly conserved families of structural patterns, called pathogen associated molecular patterns (PAMP's), which are integral parts of pathogens, and are perceived as danger signals by the host. Examples of PAMP's include bacterial cell wall structures that are absent from the host such as lipopolysaccharide (LPS) of Gram-negative bacteria, lipoteichoic acid, mannans, DNA sequences containing unmethylated CpG dinucleotides (CpG DNA), flagellin, and peptidoglycan (PGN) (196-197). The recognition of PAMP's is mediated by sets of highly conserved pattern recognition receptors (199), each of which binds to a variety of PAMP's. Cellular activation by these receptors results in acute inflammatory responses that include the production of a diverse set of cytokines and chemokines, direct local attack against the invading pathogen, and the initiation of responses that activate and regulate the adaptive component of the immune response.

The discovery of toll-like receptors (TLRs) less than a decade ago has advanced our understanding of the early events in microbial recognition and response and the subsequent development of an adaptive immune response (200-206). The Toll protein was first discovered in *Drosophila*, in which it has a pivotal role in embryonic development and microbial detection. Subsequently, a family of proteins structurally related to Toll was identified in higher organisms. Collectively, these transmembrane receptor proteins are referred to as TLRs. To date, eleven members of the mammalian TLR family have been identified, each potentially recognizing a discrete class of PAMP (207). For example, lipopolysaccharides (LPS) are recognized by TLR4, bacterial flagellin by TLR5, double-stranded RNA by TLR3 (208), and bacterial DNA by TLR9. The most recently discovered member of this family, TLR11, plays a critical role in the recognition and control of uropathogenic bacteria, and two recent studies have demonstrated that TLR3 is involved in the recognition of single-stranded viral RNA. Although it was initially believed that TLR2 in combination with TLR1 or TLR6 recognizes peptidoglycan (PGN), recent studies with highly purified PGN indicate otherwise (209). Instead, it appears that NOD proteins (NOD1 and NOD2) (210-211) and peptidoglycan recognition proteins (PGRPs) (212) are the pattern recognition receptors that detect PGN.

PGRPs are a relatively new class of pattern recognition receptors that are highly conserved from insects to mammals (213-215). *Drosophila* has 13 PGRP genes that are transcribed into at least 17 PGRPs (215). These PGRPs can be divided in extracellular (*e.g.* PGRP-SA), transmembrane (*e.g.* PGRP-LC), and intracellular or secreted (*e.g.* PGRP-LE) proteins (212). To date, 4 PGRPs have been discovered in humans, namely PGRP-S, PGRP-Iα, PGRP-Iβ and PGRP-L (213-214).

The different PGRPs may exhibit a selectivity for PGN derived from a particular group of microbes. In this respect, PGNs are large polymers composed of alternating β(1-4)-linked *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) residues, cross-linked by short peptide bridges (Figure 3.1) (216). Depending on the amino acid composition of position 3 of the peptide chain, PGNs are classified as either L-lysine-type (Lys-type) or *meso*-diaminopimelic acid-type (Dap-type). The lysine-type, typical for Gram-positive bacteria, is normally connected to the D-Ala of another peptide chain by a short bridge varying in length and amino acid composition, depending on the bacteria. In the case of Gram-negative bacteria and Gram-positive bacilli, *m*-diaminopimelic acid (Dap-type) is normally found as the third amino acid and is directly connected to D-Ala of another peptide chain.



General structure of PGN

R^2 = H or Cross-linking

R^1 = H (Lys) Gram positive PGN

R^1 = COOH (Dap) Gram negative PGN

1. Ac-L-Ala-D-isoGlu-Lys-D-Ala-D-Ala
2. Ac-L-Ala-D-isoGlu-Dap-D-Ala-D-Ala

3. R = L-Ala-D-isoGlu
4. R = L-Ala-D-isoGlu-L-Lys
5. R = L-Ala-D-isoGlu-L-Lys-D-Ala
6. R = L-Ala-D-isoGlu-L-Lys-D-Ala-D-Ala
7. R = L-Ala-D-isoGlu-Dap
8. R = L-Ala-D-isoGlu-Dap-D-Ala
9. R = L-Ala-D-isoGlu-Dap-D-Ala-D-Ala

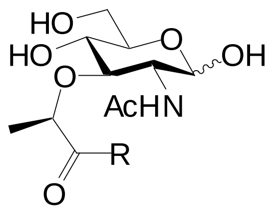


Figure 3.1: Structure of peptidoglycan (PGN) and synthetic compounds.

Drosophila PGRP-SA has been shown to interact with lysine-type PGN, activating the Toll receptor pathway (217). On the other hand, PGRP-LC and PGRP-LE recognize Dap-type PGN activating the *Imd*/*Relish* pathway (218-222). PGRPs have high homology with the T7 lysozyme, a type 2 *N*-acetylmuramoyl-L-alanine amidase that hydrolyses the bond between MurNAc and L-Ala of PGN (223). In this respect, *Drosophila* PGRP-SC1b and PGRP-LB have been shown to possess amidase activity

(224).

Each of the four mammalian PGRPs (PGRP-L, PGRP-Iα, PGRP-Iβ and PGRP-S) is able to bind peptidoglycan, however, possible selectivities for lysine or Dap-type PGN have either not been determined or remain controversial. In addition, the mode of cellular activation and bactericidal activity of these PGRP's is largely unknown. The limited data for PGRP-L indicates that this protein exhibits lytic activity (225-226). The function of PGRP-Iα and PGRP-Iβ is unknown and most research has thus far focused on PGRP-S. Mouse PGRP-S found in neutrophil tertiary granules participates in the intracellular neutralization of bacteria. Mice deficient in this PGRP are much more susceptible to intraperitoneal infections with low-pathogenic Gram-positive bacteria (227). However, bovine PGRP-S, located in neutrophil and eosinophil granules has been shown to inhibit the growth of both Gram-positive and -negative bacteria (228).

To determine in the ligand requirements for various PGRP's, we have synthesized a range of partial structures of PGN (Figure 3.1) that contain lysine or DAP as the third amino acid. The interactions of these compounds with human PGRP-S and the C-terminal domain of human PGRP-Iα containing two tandem domains, have been studied in real time using surface plasmon resonance (SPR).

3.3 Experimental Procedures

3.3.1 Materials

Procedures for the expressing recombinant PGRP-IαC (residues 177-341) by *in-vitro* folding from *Escherichia coli* inclusion bodies have been described previously (229).

3.3.2 Chemical Synthesis of PGN part structures

Compound **6**: Sieber Amide resin (230) (100 mg, 42 μmol, Novabiochem) was swelled in dry dimethylformamide (DMF, ~ 120 min, 3 ml), treated with 20% piperidine in DMF (3 x 5 min, 3 x 2 ml), washed with freshly distilled DMF (3 x 3 ml), and then reacted with Fmoc-D-Ala-OH (26.12 mg, 84 μmol, Novabiochem) in DMF by using PyBOP (43.7 mg, Novabiochem), 1-hydroxybenzotriazole (HOBt) (11 mg, Aldrich), and *N,N*-diisopropylethylamine (DIPEA) (29.2 μl, Alfa Aesar, Ward Hill, MA). Progress of the reaction was monitored by the Kaiser test. After completion of the coupling, the resin was washed with (3 x 3 ml) and the Fmoc protecting group was removed with 20% piperidine in DMF (3 x 5 min, 3 x 2 ml). The reaction cycle was repeated using Fmoc-D-Ala-OH (26.1 mg, 84 μmol), Fmoc-L-Lys(Mtt)-OH (52.4 mg, 84 μmol), Fmoc-D-isoglutamine (30.9 mg, 84 μmol), Fmoc-L-Ala-OH (26.12 mg, 84 μmol, Novabiochem), and, subsequently, 2-*N*-acetyl-1-β -*O*-allyl-4,6-benzylidene-3-muramic acid (231) (35.4 mg, 84 μmol). The resulting resin-bound glycopeptide was washed with DMF (3 x 3 ml), dichloromethane (7 x 3 ml), and methanol (3 x 3 ml). The resin was dried *in vacuo* for 4 hrs and reswelled in dichloromethane (DCM) (~5 ml) and filtered. The glycopeptide was released by treatment of the resin with 2% trifluoroacetic acid (TFA) in DCM (10 x 2 ml). The combined washings were concentrated under reduced pressure and co-evaporated with toluene (3 x 10 ml) to remove traces of trifluoricacetic acid (TFA). The crude product was subjected to 20% TFA in DCM to ensure complete removal of the benzylidene protecting group. The resulting product was purified by Sephadex G15 size exclusion column (Amersham Biosciences) chromatography to give [allyl-2-*N*-acetyl-3-*O*-muramyl]-L-alanyl-D-isoglutamyl-L-lysine (23.4 mg, 70%) as a white amorphous solid. ^{1}H NMR (500 MHz, D_2O): δ = 5.87-5.93 (1H, m, OCH_2CHCH_2), 5.25-5.32 (2H, dd, OCH_2CHCH_2), 4.53 (1H, d, *H*1, *J* = 8.3 Hz), 4.13-4.35 (8H, m, α *H*-ala x 3, α *H*-lys, α *H*-glu, α *H*-lactic acid, OCH_2CHCH_2), 3.93 (1H, d, *H*6a, *J* = 12.2 Hz), 3.85 (1H, t, *H*2), 3.76-3.79 (1H, dd, *H*6b), 3.45-3.57 (3H, m, *H*3, *H*4, *H*5), 3.00 (2 H, t, ε CH_2-lys), 2.36-2.43 (2H, m, γ CH_2-glu), 2.12-2.18 (1H, m, β CH_2-glu), 1.94-2.03 (4H, m, β CH_2-glu, NH*Ac*), 1.67-1.82 (4H, m, β, δ-CH_2-lys), 1.37-1.45 (14H, m, CH_3-lactic acid, γ CH_2-lys,CH_3-ala x 3) ^{13}C NMR (75 MHz, D_2O): 177.88, 175.98, 175.93, 175.31, 175.22, 174.92, 174.83, 174.26, 133.52 (OCH_2CHCH_2), 118.23 (OCH_2CHCH_2), 100.27 (C1), 82.98, 75.78, 68.84 (C3, C4, C5), 70.66 (OCH_2CHCH_2), 60.86 (C6), 55.29, 54.30, 52.89, 50.14, 49.95, 49.60 (α-Cs), 39.30 (ε CH_2-lys), 31.46 (γ CH_2-glu), 30.25 (β CH_2-glu), 26.51 (β CH_2-lys), 22.34, 22.21 (δ CH_2-lys, $NHCH_3$), 18.91, 16.68, 16.36. HRMS-MALDI-TOF calcd for $C_{34}H_{59}N_9O_{13}$ (M+Na): 824.4232, found 824.3087. The compound (10.6 mg, 12.4 μmol) was dissolved in a mixture of ethanol/acetic acid/water (EtOH/HOAc/H_2O, 2:1:1, 0.8 ml), and 10% Pd on charcoal (9 mg) was added. After stirring at room temperature for 48 h, the reaction mixture was filtered. The filtrate was concentrated under reduced pressure, and the residue was coevaporated from toluene (3 x 20 ml). The residue was subjected to Sephadex G15 size exclusion column chromatography to give the target compound **6** as a mixture of α /β anomers (8.6 mg, 91%). ^{1}H NMR (500 MHz, D_2O): δ 5.04 (0.60H, d, *H1*-α anomer, *J* = 3.3 Hz), 4.56 (0.39H, d, *H*-1-β -anomer, *J* = 8.4 Hz), 4.17-4.08 (6H, m, α *H*-Lys, α *H*-Glu, α *H*-Ala x 3, α *H*-3-propionic acid), 3.36-3.86 (6H, m, *H2,H3,H4,H5,H6*), 2.87 (2H, t, ε CH_2-lys), 2.21-2.29 (2H, m, γ CH_2-glu), 2.19-2.03 (1H, m, β CH_2-glu), 1.82-1.87 (4H, m, CH_2-glu, NH*Ac*), 1.54-1.67 (4H, β, δ–CH_2-lys), 1.37-1.45 (14H, m, γ CH_2-lys, CH_3-lactic acid, CH_3-ala x 3). ^{13}C NMR (75 MHz, D_2O) 177.89, 176.14, 175.94, 175.32, 175.21, 174.93, 174.82, 174.38, 174.13, 95.07 (C1-α), 91.13 (C1-β) , 82.78, 79.87, 78.24, 77.90, 75.87, 71.65, 69.03, 68.81, 60.87, 60.68, 56.33, 54.30, 53.86, 52.89, 50.13, 49.96, 49.61, 39.30 (ε *C*H_2-lys), 31.42 (γ *C*H_2-glu), 30.26, 27.08, 26.51, 22.38, 22.22, 22.15, 18.81, 16.73, 16.68, 16.37. HRMS- MALDI-TOF calcd for $C_{31}H_{55}N_9O_{13}$ (M + Na): 784.8213, found 784.5895.

Compounds **5-6** were synthesized using similar protocol whereby suitable amino acids were chosen depending on the desired target. Analytical data for the glycopeptides is listed below.

MTrP_Lys (**5**): Yield 47%, ^{1}H NMR (500 MHz, D_2O): δ = 5.04 (0.45H, d, *H*1-α anomer, *J* = 3.5 Hz), 4.55 (0.54H, d, *H*-1-β -anomer, *J* = 8.0 Hz), 4.10-4.17 (5H, m, α *H*-Lys, α *H*-Glu, α *H*-Ala x 2, α -*H*-lactic acid), 3.34-3.81 (6H, m, H*2*,H*3*,H*4*,H*5*,H*6*), 2.88 (2H, t, ε CH_2-lys), 2.25-2.28 (2H, m, γ CH_2-glu), 2.10 (1H, m, β CH_2-glu), 1.82-1.91 (4H, m, CH_2-glu, NH*Ac*), 1.54-1.72 (4H, β, δ−CH_2-lys), 1.25-1.37 (11H, m, γ CH_2-lys, CH_3-lactic acid, CH_3-ala x 2). ^{13}C NMR (HSQC, D_2O) 95.07 (C1-α), 91.47(C1-β), 82.99, 80.06, 78.20, 75.81, 73.92, 71.95, 69.29, 60.91, 54.39, 53.99, 52.93, 39.34 (ε *C*H_2-lys),, 31.63 (γ *C*H_2-glu),, 30.83, 27.10, 26.97, 26.57, 22.98, 22.19, 18.93, 16.71, 16.41. HRMS-MALDI-TOF calcd for $C_{28}H_{50}N_8O_{12}$ (M + Na): 713.3548, found 713.4080.

MTP_Lys (**4**): Yield 61%, ^{1}H NMR (500 MHz, D_2O): δ 5.16 (0.69H, d, *H*1-α anomer, *J* = 3.3 Hz), 4.67 (0.31H, d, *H*1-β -anomer, *J* = 8.1 Hz), 4.20-4.34 (4H, m, α-*H*, lys, α-*H*, glu, α-*H*, ala, α-*H*, lactic acid), 3.50-4.00 (6H, m, *H*2,*H*3,*H*4,*H*5,*H*6), 3.01 (2H, t, ε-CH_2, lys), 2.39-2.45 (2H, m, γ-CH_2, glu), 2.15-2.23 (1H, m, β-CH_2, glu), 1.65-2.00 (8H, m, β-CH_2, glu, β, δ-CH_2, lys, NHCOCH_3), 1.37-1.47 (8H, m, γ-CH_2, lys, CH_3, lactic acid, CH_3, ala). ^{13}C NMR (75 MHz, D_2O) 177.06,176.02,175.25, 174.14, 95.07 (C1-β), 91.13 (C1-α), 82.77, 79.85, 78.46, 78.24, 77.89, 75.87, 73.51, 71.66, 69.05, 60.88, 60.67, 56.34, 53.86, 53.63, 52.87, 49.98, 49.03, 39.33 (ε-CH_2, lys), 31.57 (γ-CH_2, glu), 30.53, 26.93, 26.42, 22.65, 22.37, 22.28, 22.15, 18.79, 16.70. HRMS- MALDI-TOF calcd for $C_{25}H_{45}N_7O_{11}$ (M + Na): 642.3067, found 642.3777.

The DAP containing muramyl tripeptides **7-9** were synthesized similarly to procedures described above whereby the Lys(Mtt)-OH was replaced by a suitable protected DAP derivative (232) (38.7 mg, 42 μmol) to afford the desired protected DAP derivative. Once cleaved from the resin, the DAP-PGNs were treated with 20% TFA to deprotect the tert Butoxycarbonyl (Boc) and tert butyl (t-Bu) protecting groups on the side chain of the DAP. The deprotected derivative was precipitated from cold diethyl ether to afford an off white compound. To a solution of this compound in EtOH: H_2O: 1(N) HCl (4:2:0.01, 0.6ml), 10% Pd/C (5 mg) was added and stirred at room temperature for 16h. The solution was filtered and purified by Sephadex G10 size exclution column chromatography to afford the target compound as a mixture of α /β anomers (9.3 mg, 30% overall).

MTP_DAP (**7**): ^1H NMR (500MHz, D_2O): δ 5.08 (0.16H, d, *H*1-α anomer), 4.32 (0.84H, d, *H*1-β anomer, *J* = 8.4Hz), 4.07-4.23 (5H, m, α-*H* x 2**,** DAP, α-*H*, ala, α-*H*, glu, α-*H*, lac), 3.67-3.89 (3H, m, *H*2, *H*6ab), 3.38-3.50 (3H, m, *H*3, *H*4, *H*5), 2.27-2.36 (2H, m, γ-CH_2, glu), 2.10 (1H, m, β-C*H*H, glu), 1.65-1.93 (8H, m, β,δ-CH_2, DAP, β-C*H*H, glu, NHCOCH_3), 1.42-1.45 (2H, m, γ-CH_2, DAP), 1.29-1.35 (6H, m, CH_3, Lac, CH_3, ala). ^{13}C (HSQC): 102.18 (*C*-1-β), 91.74 (*C*-1-α), 83.31, 78.93, 76.24, 69.16, 60.66 (*C*6), 60.65, 60.06, 58.04, 55.68, 55.01, 54.33, 53.66, 50.29, 32.09 (γ-*C*, glu), 30.75 (*C*-Dap), 27.71 (β-*C*, glu), 23.00 (NHCO*C*H$_3$), 21.98, 19.29, 17.27. HRMS- MALDI-TOF calcd for $C_{25}H_{45}N_7O_{11}$ (M + HCl): 700.1355, found 700.4058.

MTrP_DAP (**8**): ^{1}H NMR (600MHz, D_2O): δ 5.19 (0.16H, bs, *H*1), 4.48 (0.84H, d, *H*1, *J* = 7.2Hz), 4.23-4.39 (6H, m, α-*H*, ε-*H*, DAP, α-*H* x 2, ala, α-*H*, glu, α-*H*, lac), 3.51-3.98 (6H, m, *H*2, *H*6, *H*3, *H*4, *H*5), 2.40 (2H, m, γ-C*H*$_2$, glu), 2.13-2.14 (1H, m, β-CH*H*, glu), 1.78-2.09 (8H, m, β,δ-C*H*$_2$, DAP, β-C*H*H, glu, NHCOC*H*$_3$), 1.38-1.50 (11H, m, γ-C*H*$_2$, DAP, C*H*$_3$, Lac, C*H*$_3$ x 2, ala). ^{13}C (HSQC): 102.13 (*C*-1-β), 91.65 (*C*-1-α), 81.82, 78.70, 77.49, 76.28, 72.13, 69.36, 61.22 (*C*6), 60.01, 59.15, 57.59, 57.42, 55.68, 54.79, 54.30, 53.13, 53.95, 49.97, 31.78 (γ-*C*, glu), 30.87 (*C*-Dap), 27.81 (β-*C*, glu), 22.92 (NHCO*C*H$_3$), 21.24, 19.25, 18.95, 18.03, 16.81. HRMS- MALDI-TOF calcd for $C_{25}H_{45}N_7O_{11}$ (M + HCl): 771.2135, found 771.8770.

MPP_DAP (**9**): ^{1}H NMR (500MHz, D_2O): δ 5.16 (0.41H, bs, *H*1), 4.40 (0.58 H, d, *H*1, *J* = 9.0 Hz), 4.22-4.29 (7H, m, α-*H*, ε-*H*, DAP, α-*H* x 3, ala, α-*H*, glu, α-*H*, lac), 3.50-3.98 (6H, m, *H*2, *H*6, *H*3, *H*4, *H*5), 2.39 (2H, m, γ-C*H*$_2$, glu), 2.13-2.15 (1H, m, β-CH*H*. glu), 1.90-2.03 (8H, m, β,δ-C*H*$_2$, DAP, β-C*H*H, glu, NHCOC*H*$_3$), 1.38-1.57 (14H, m, γ-C*H*$_2$, DAP, C*H*$_3$, Lac, C*H*$_3$ x 3, ala). ^{13}C (HSQC): 102.39 (*C*-1-β), 91.66 (*C*-1-α), 83.35, 80.13, 78.52, 78.25, 76.11, 72.08, 69.40, 60.95 (*C*6), 60.88, 60.60, 57.40, 55.59, 54.61, 54.40, 53.77, 53.01, 50.01, 31.60, 30.48 (*C*-Dap), 27.05, 22.40 (NHCO*C*H$_3$), 21.01, 19.06, 16.83. HRMS- MALDI-TOF calcd for $C_{25}H_{45}N_7O_{11}$ (M + HCl): 842.2914, found 842.6710.

3.3.3 SPR analysis of PGRP- ligand interactions

The biospecific interaction analysis was performed using BIAcore 3000 biosensor system (Biacore Inc. Uppsala Sweden). The CM-5 research grade sensor chip, HBS-EP buffer, and immobilization reagents (1-ethyl-3(3-*N*,*N*,-dimethylaminopropyl)carbodiimide (EDC), *N*-hydroxy succinimide (NHS), and ethanolamine) were obtained from Biacore Inc. Uppsala Sweden, phosphate-buffered saline (PBS) buffer was purchased from Sigma and MDP was purchased from CalBioChem. All solutions were filtered using a 0.22 μm PES membrane syringe filter and degassed prior to use. PGRP-IαC was covalently immobilized by a standard amine coupling procedure using the amine coupling kit supplied by the manufacturer. A fixed flow rate of 10 μl/min was used throughout the immobilization procedure with HBS-EP (pH 7.4, 0.01M HEPES, 150mM NaCl, 3mM EDTA, 0.005% P20) as the running buffer. The surface was activated using 70 μl of freshly mixed 1:1 100 mM NHS and 391 mM EDC for 7 minutes. Upon activation, a 60 μg/ml solution of PGRP-Iα in 10 mM NaOAc (pH 4.5) was injected for 8 minutes. The remaining active esters on the surface were quenched using 70 μl of 1.0 M ethanolamine (pH 8.5) for 7 minutes. A ligand density of approximately 10,000 RU was achieved. Blocking of a control flow cell was accomplished by activation followed by an immediate quenching of the flow cell as described above. PGRP-S immobilization was accomplished using the same protocol as above with a change to 5 mM maleate buffer (pH 6.0) as the immobilization buffer. A ligand density of approximately 5,000 RU was achieved after 8 minutes of protein injection. Although a higher initial immobilization density was targeted, longer PGRP-S injections resulted in no change of protein density. Initial binding studies of the DAP containing compounds with PGRP-S suggested high affinity prompting the formation of a low density immobilization surface for kinetic analysis. A similar protocol was used with an injection of PGRP-S solution of 50 μg/ml in 10 mM NaOAc (pH 4.5) for 5 minutes to afford an immobilized surface of 2700 RU.

For the binding studies, fixed flow rates of 5 μl/min for association and dissociation with a constant temperature of 25°C were employed. The association and dissociation times were 5 minutes and 10 minutes, respectively. These surfaces had greater than 90% reproducibility if used within a 3-4 day period. PGRP Iα: phosphate-buffered saline (PBS) buffer (pH 7.4, 0.01 M phosphate, 138 mM NaCl, 2.7 mM KCl) was selected as both the running and dissociation buffer. The lysine containing analytes, pentapeptide (Lys-PP **1**), muramyldipeptide (MDP **3**), muramyltripeptide (MTP **4**), muramyltetrapeptide (MTrP **5**) and muramylpentapeptide (MPP **6**), were passed over the surface at a concentration range from 20 to 60 μM using the kinetic wizard method. The surface did not require a regeneration injection for the Lys-PP, MDP, or MTP. For MTrP and MPP, a regeneration injection of 0.01% Tween-20 in water for 20 seconds at a flow rate of 30 μl/min followed by 15 minutes of stabilization time was required to achieve prior baseline status. The DAP containing analytes, DAP-PP (**2**), MTP-DAP (**7**), MTrP-DAP (**8**) and MPP-DAP (**9**), were passed over the surface at concentrations ranging from 10 to 50 μM. For MTrP-DAP and MPP-DAP, the regeneration was performed as described for the lysine-containing fragments.

For PGRP-S, HBS-EP buffer was selected as both the running and dissociation buffers. Due to a greater variation in binding affinity, optimum concentration ranges for each analyte were established. For the lysine containing compounds, MTP was examined from 500 to 1000 μM, MTrP from 200 to 500 μM, and MPP from 50 to 300 μM. Initial binding studies were conducted on a flow channel with 5000 RU density. The DAP containing muramyl peptides showed high affinity biding with *on* and *off* rates in the measurable range for kinetic analysis. These studies were conducted on a flow channel containing 2700 RU of PGRP-S to minimize rebinding and mass transport effect. A concentration range of 10 to 1000 nM was selected for MPP-DAP and MTrP-DAP and for MTP-DAP, a concentration range of 100 nM to 1000 nM was selected. The kinetic studies were performed using wizard kinetic software with a flow rate of 10 µl/min and association and dissociation times of 5 and 10 minutes, respectively. The surface was regenerated by a 60 second injection of 10 mM NaOH, pH 11.4, at a flow rate of 30 µl/min.

3.3.4 Data analysis

The responses near equilibrium (Req) for the comparative affinity and rate constants k_1 and k_{-1} for kinetic analysis were obtained by fitting the primary sensogram data using the BIA evaluation 3.1 software. The dissociation rate constant is derived using:

$$R_t = R_{t_0}\, e^{-k_{-1}(t - t_0)} \qquad \{3.1\}$$

where R_{t_0} is the amplitude of the initial response, and k_{-1} is the dissociation rate constant. The association rate constant k_1 can be derived from the measured k_{-1} values, using:

$$R_t = R_{max} [1 - e^{-(k_1 C + k_{-1})(t - t_0)}] \qquad \{3.2\}$$

where R_t is response at time t, R_{max} is the maximum response, C is concentration of the analyte in the solution, and k_1 and k_{-1} are association and dissociation rate constants, respectively. The ratio of k_1 and k_{-1} yields the value of association constant K_A (k_1/k_{-1}).

3.4 Results

The interactions of human PGRP-IαC and PGRP-S with a range of synthetic lysine- and DAP-containing PGN fragments were probed using surface plasmon resonance (SPR). SPR is a rapid and sensitive method for the evaluation of affinities of bimolecular interactions (233). A benefit of this method is that it relies exclusively on mass changes, thus allowing interactions to be studied in real-time without the need for external labels such as fluorophores, which in some cases can alter the nature of the interaction. In this study, recombinant human PGRP-Iα and PGRP-S were immobilized on research-grade CM5 sensor chips and the synthetic compounds **1-9** were employed as analytes. Collecting SPR data for low molecular weight analytes, such as compounds **1-9**, is challenging because the refractive index monitored during a binding event is relatively small, thus resulting in responses with much lower magnitudes than those observed in typical protein-protein interactions. Despite these challenges, the high sensitivity and reproducibility of modern instruments combined with proper experimental design permits the direct monitoring of the binding of low molecular weight analytes to immobilized proteins (234). An alternative approach whereby the synthetic peptidoglycan part structures would be immobilized and the PGRPs employed as analytes is expected to lead to artifacts because the immobilization may destruct vital functional groups of the small synthetic compounds.

The compounds **1**-**9** (Figure 3.1) were synthesized by polymeric support synthesis using a Sieber Amide resin, Fmoc-protected amino acids and a properly protected muramic acid derivative. The synthetic compounds were designed in such a manner that the significance of each amino acid and the muramic acid moiety could be addressed for binding. Furthermore, by employing two series of compounds that contain either lysine or DAP as the third amino acid, the importance of these residues could be studied.

A relatively high immobilization of approximately 10,000 RU of PGRP-IαC was accomplished on the *N*-hydroxy succinimide (NHS)-activated groups of a CM-5 research grade sensor chip surface to achieve good signal to noise ratios. For all SPR studies, bulk refraction caused by the difference in refractive index of the running buffer and sample injection was negated by using a control cell that was functionalized by ethanolamine.

A representative sensogram for the interaction of the lysine-containing muramyl pentapeptide **6** with the immobilized PGRPI-αC is shown in Figure 3.2. Fitting using

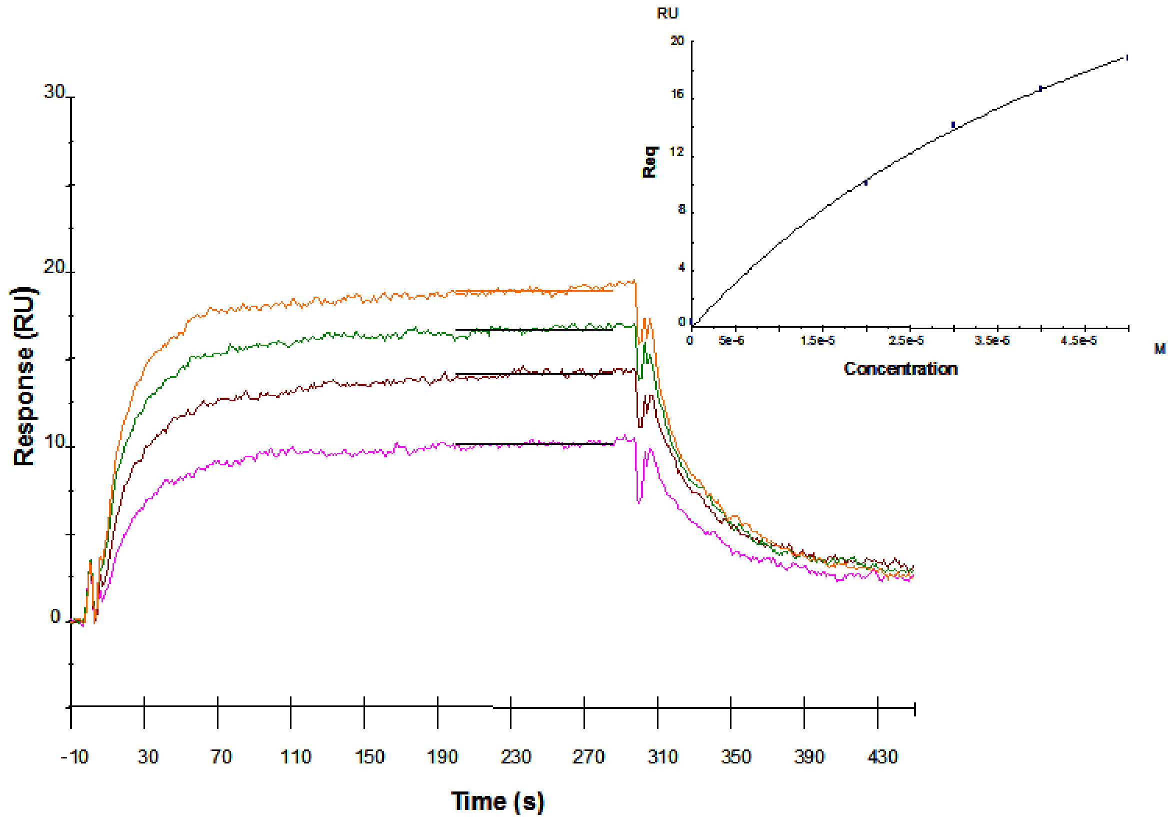


Figure 3.2: Sensorgram representing the concentration dependent binding of Lysine containing MPP **(6)** with immobilized PGRP-Iα (10,000 RU). Steady-state binding analysis of MPP-Lys (**6**) at concentrations from bottom to top of 20, 30, 40 and 50 μM with PGRP-Iα resulted in a K_D of 6.2 x 10^{-5} M. *Inset plot* shows non-linear steady-state affinity analysis.

steady state conditions gave a dissociation constant K_D = 62 μM (Table 3.1). The compounds **4-5** showed binding at concentrations greater than 50 μM, whereas peptide **1** (Lys-PP) (Figure 3.3) and MDP (**3**) (Figure 3.4) exhibited no binding at concentrations up to 1 mM. The latter results indicate that the muramyl moiety is required for

Table 3.1. Binding constants (K_D) of PGN fragments with human PGRPs

Analyte	PGRP-Iα	PGRP-S
MPP-Lys (**6**)	62 μM	189 μM
MTP-Dap (**7**)	NA*	326 nM
MTrP-Dap (**8**)	NA*	92.4 nM
MPP-Dap (**9**)	NA*	104 nM

*Due to biphasic interactions at higher concentrations, kinetic values could not be determined.

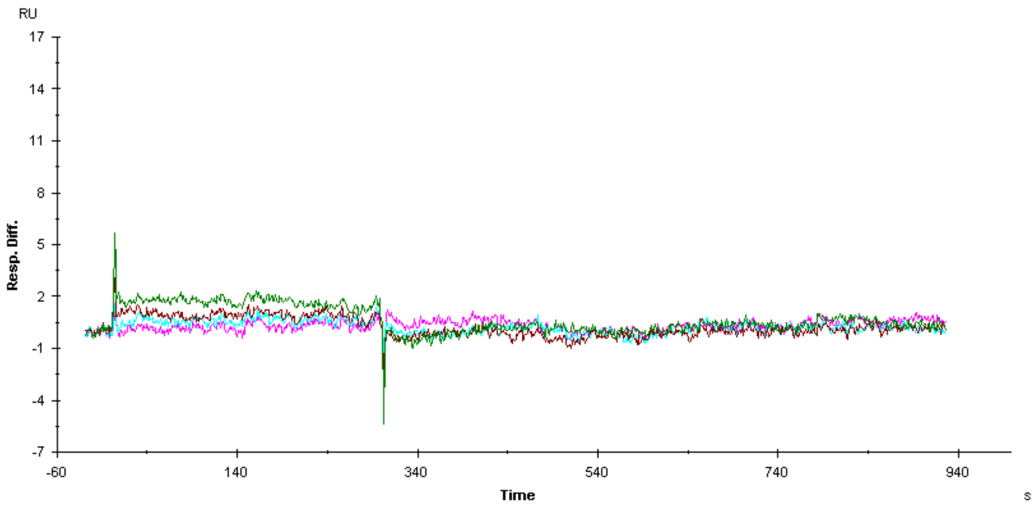


Figure 3.3: Sensorgram of LPP (**1**) at concentrations up to 1mM showing no detectable binding with PGRP-Iα.

complexation and that MTP **4** is the smallest fragment of PGN recognized by PGRP-Iα. Equilibrium based kinetic analysis was not possible for MTP **4** and MTrP **5** due to an

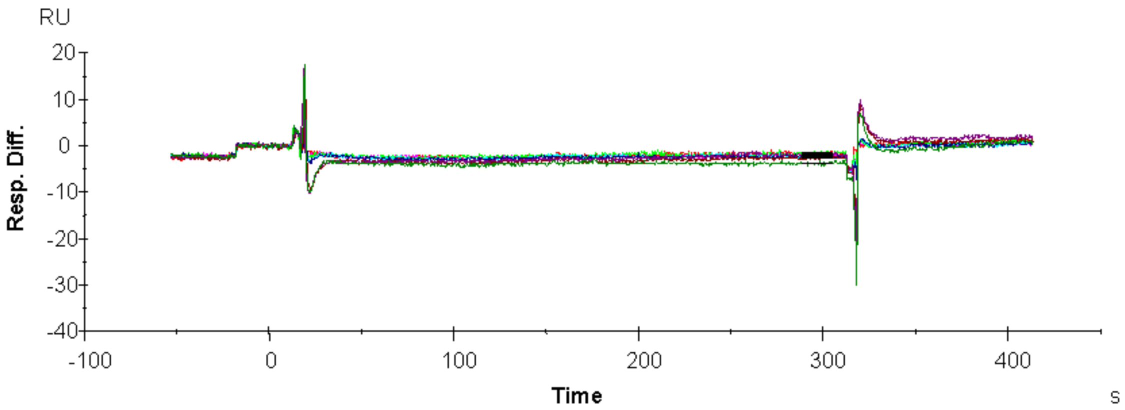


Figure 3.4: Sensorgram of MDP (**3**) at concentrations up to 1mM showing no detectable binding with PGRP-Iα.

apparent biphasic nature of the real-time sensogram slope shifts at higher concentrations (Figures 3.5 and 3.6). Given that SPR detection is based on changes in mass at the surface of a sensor chip, normalization of respective RUs by the molecular weight of

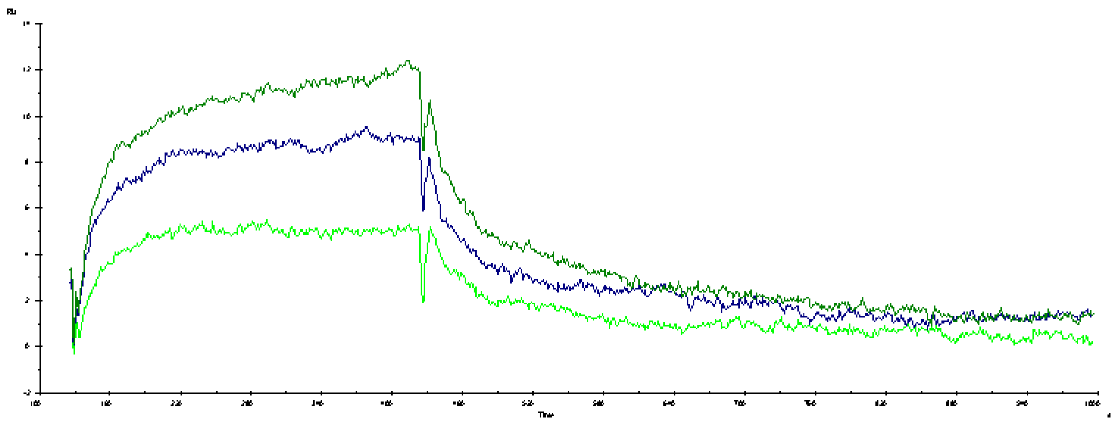


Figure 3.5: Sensorgram representing the concentration dependent binding of Lysine containing MTP (**4**) with immobilized PGRP-Iα (10,000 RU). Concentrations from bottom to top of 40, 50 and 60μM were injected over the surface.

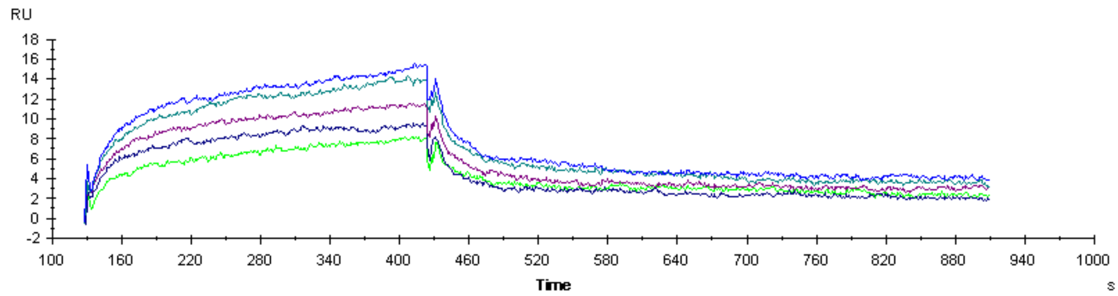


Figure 3.6: Sensorgram representing the concentration dependent binding of Lysine containing MTrP (**5**)with immobilized PGRP-Iα (10,000 RU). Concentrations from bottom to top of 20, 30, 40, 50 and 60μM were injected over the surface.

each analyte allows a qualitative comparison of analyte sets. A prerequisite of such a comparison is that all variables are kept constant, most importantly the protein surface employed. Normalization of the binding data of **4-6** (Figure 3.7) demonstrated that MTrP **5** and MPP **6** are significantly more potent ligands than MTP **3**. Furthermore, the data suggest that MPP **6** has a slightly higher affinity than MTrP **5**.

Normalized SPR data for the Dap-containing compounds **2** and **7-9** showed a trend similar to that of the lysine-containing derivatives **1** and **4-6**, whereby PGRP-Iα had the weakest affinity for the MTP-Dap while MTrP-Dap and MPP-Dap had similar and significantly higher affinities (Figures 3.8-12). A comparison of the normalized data for the Dap-containing compounds **7-9** and lysine-containing derivatives **4-6** indicated that the PGRP-Iα has a low selectivity for the lysine- or Dap-containing compounds.

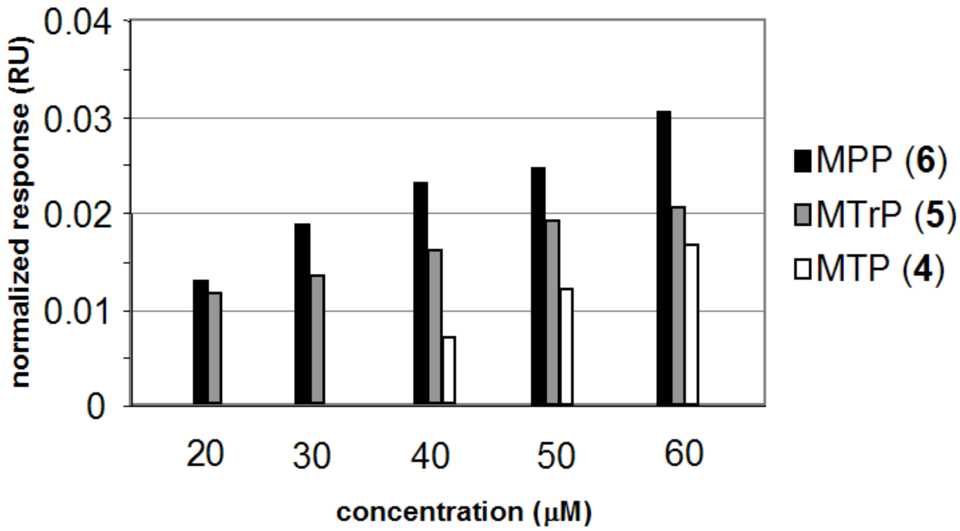


Figure 3.7: Comparative binding analysis of Lysine containing MTP **(4)**, MTrP **(5)** and MPP **(6)** with immobilized PGRP-Iα (10,000 RU). Concentrations of 20, 30, 40, 50 and 60 μM of each analyte were passed over the surface for 5 minutes. The average RU from 270 to 300 seconds was divided by respective molecular weight of each analyte to achieve a normalized RU to allow a comparison of affinity through the bar diagram. MDP (**3**) and Lysine-pentapeptide (**1**) had no detectable binding up to 1mM concentrations.

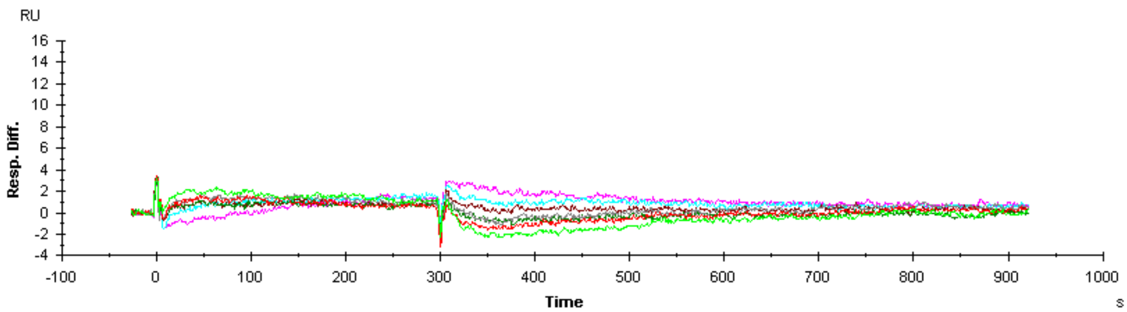


Figure 3.8: Sensorgram of DAP-PP (**2**) at concentrations of up to 1mM showing no detectable binding with PGRP-Iα.

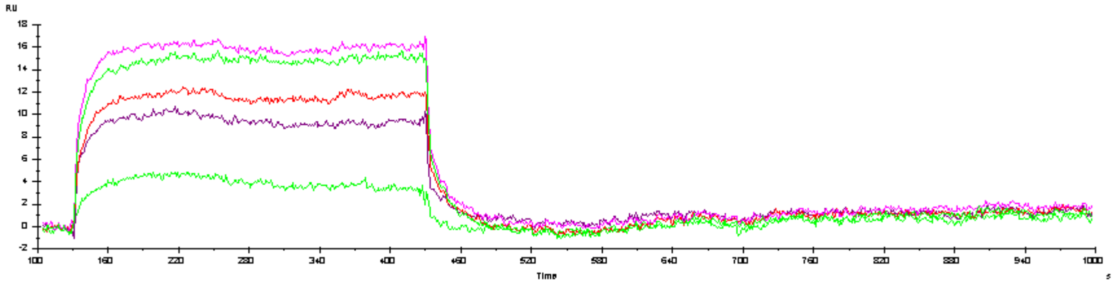


Figure 3.9: Sensorgram representing the concentration dependent binding of DAP containing MTP (**7**) with immobilized PGRP-Iα (10,000 RU). Concentrations from bottom to top of 10, 20, 30, 40 and 50μM were injected over the surface.

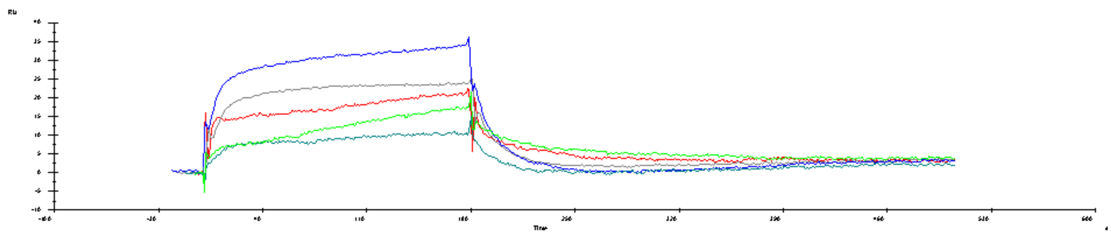


Figure 3.10: Sensorgram representing the concentration dependent binding of DAP containing MTrP (**8**) with immobilized PGRP-Iα (10,000 RU). Concentrations from bottom to top of 10, 20, 30, 40 and 50μM were injected over the surface.

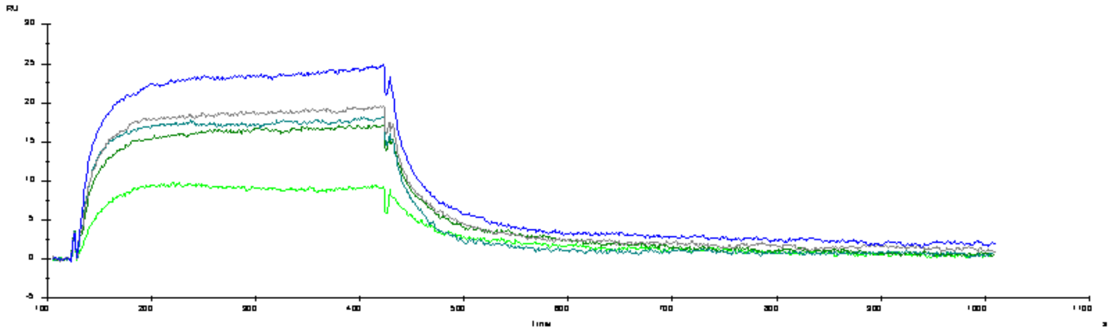


Figure 3.11: Sensorgram representing the concentration dependent binding of DAP containing MPP (**9**)with immobilized PGRP-Iα (10,000 RU). Concentrations from bottom to top of 10, 20, 30, 40 and 50μM were injected over the surface.

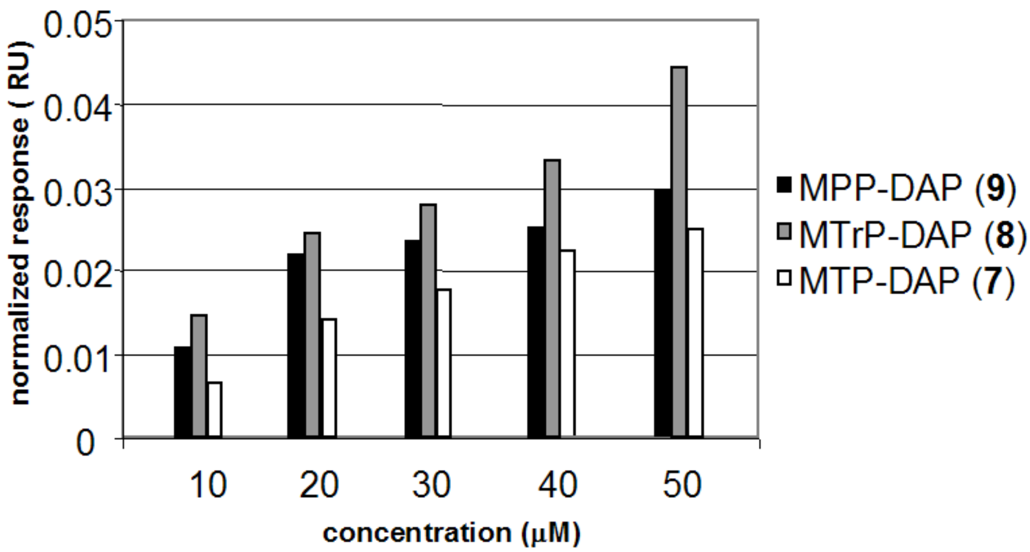


Figure 3.12: Comparative binding analysis of DAP containing MTP **(7**), MTrP **(8)** and MPP **(9)** with immobilized PGRP-Iα (10,000 RU). Concentrations of 10, 20, 30, 40 and 50 μM of each analyte were passed over the surface for 5 minutes. The average RU from 270 to 300 seconds was divided by respective molecular weight of each analyte to

achieve a normalized RU to allow a comparison of affinity through the bar diagram. MDP (**3**) and DAP-pentapeptide (**2**) had no detectable binding up to 1mM concentrations.

Next, the interactions of compounds **1-9** with a CM-5 research grade sensor chip, containing 5000 RU of PGRP-S, were studied (Figures 3.13-22). The normalized data for the binding of immobilized PGRP-S with lysine containing muramyl peptides **4-6** showed that MPP **6** has a slightly higher affinity than MtrP **5**, which are both significantly better ligands than MTP **4** (Figure 3.20). Steady-state equilibrium analysis of MPP **6** gave a K_D of 189 μM (Table 3.1).

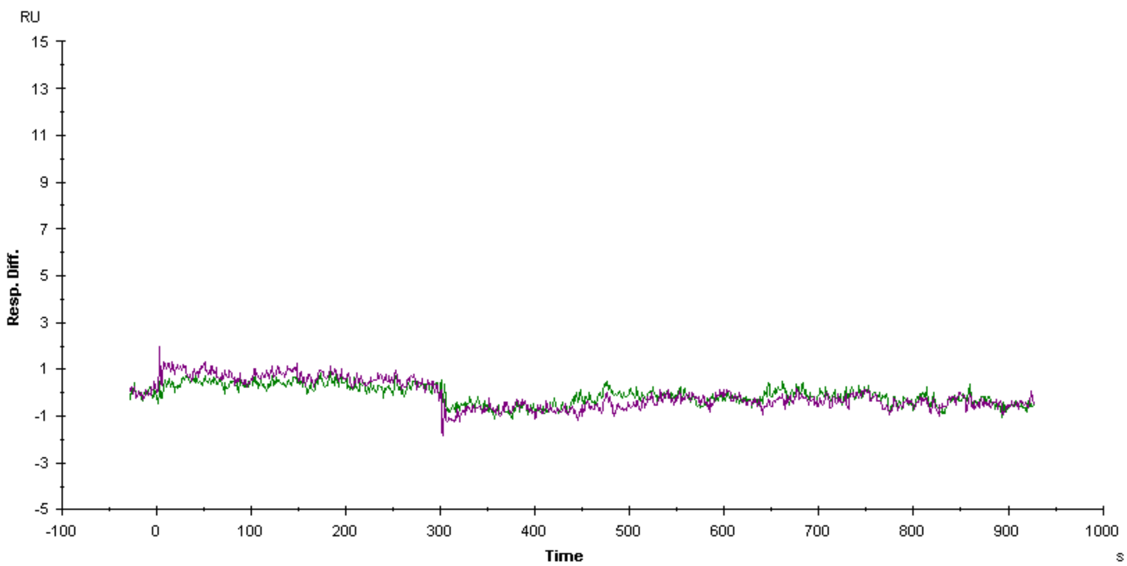


Figure 3.13: Sensorgram of LPP (**1**) at concentrations of up to 1mM showing no detectable binding with PGRP-S.

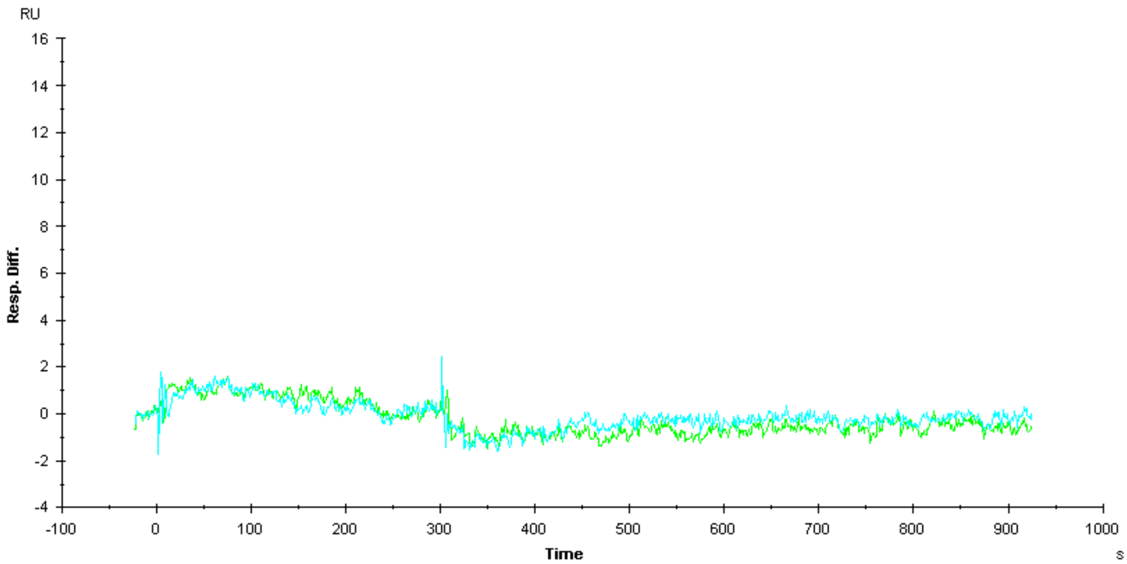


Figure 3.14: Sensorgram of DAP-PP (**2**) at concentrations of up to 1mM showing no detectable binding with PGRP-S.

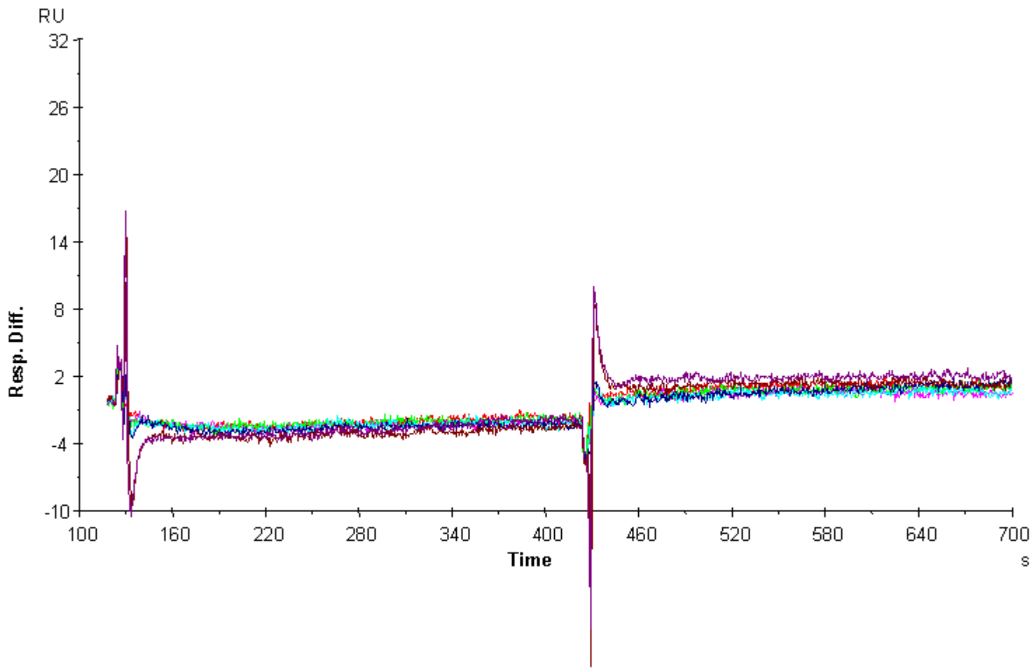


Figure 3.15: Sensorgram of MDP (**3**) at concentrations up to 1mM showing no detectable binding with PGRP-S.

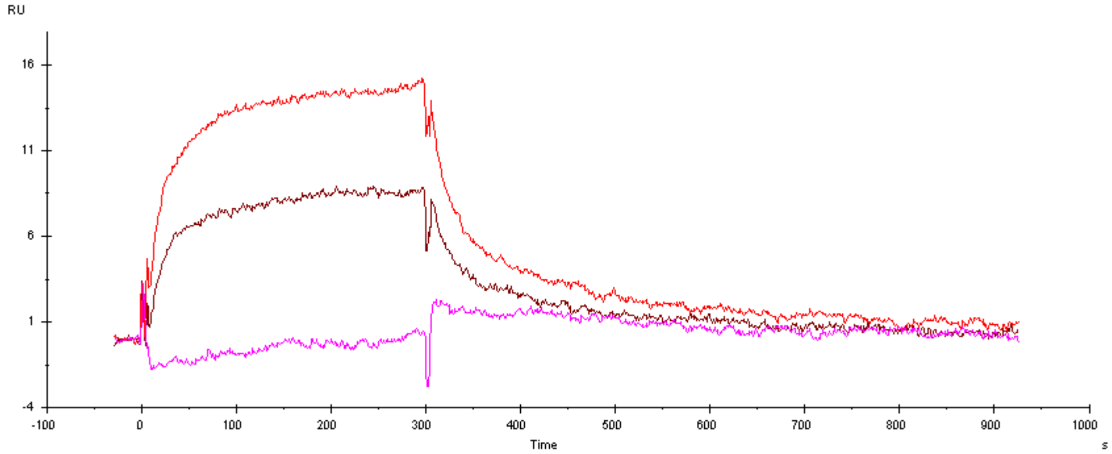


Figure 3.16: Sensorgram representing the concentration dependent binding of Lysine containing MTP (**4**)with immobilized PGRP-S (5,000 RU). Concentrations from bottom to top of 500, 800 and 1000μM were injected over the surface.

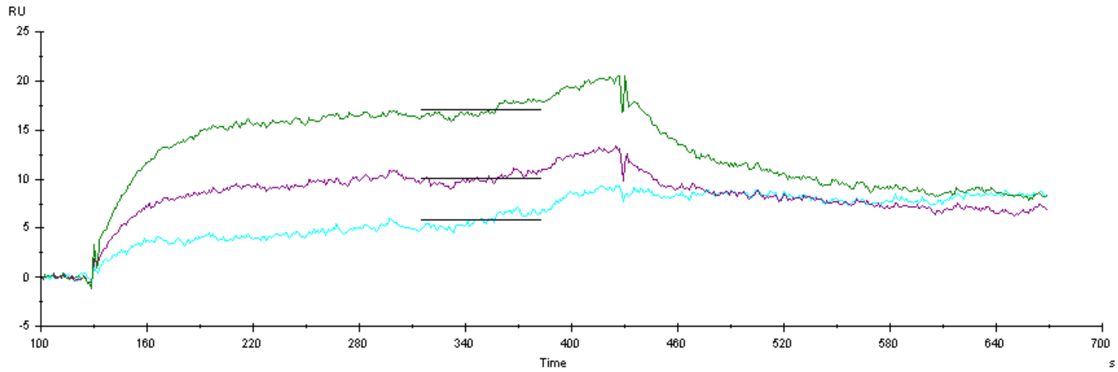


Figure 3.17: Sensorgram representing the concentration dependent binding of Lysine containing MTrP (**5**) with immobilized PGRP-S (5,000 RU). Concentrations from bottom to top of 200, 300 and 500μM were injected over the surface.

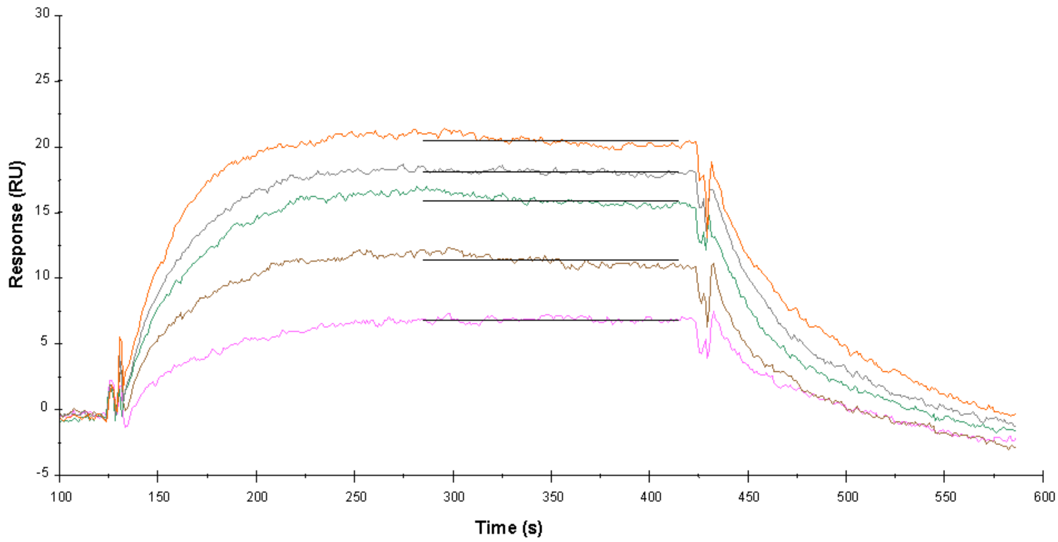


Figure 3.18: Sensorgram representing the concentration dependent binding of Lysine containing MPP with immobilized PGRP-S (5,000 RU). Steady-state binding analysis of MPP-Lys (**6**) at concentrations from bottom to top of 50, 100, 150, 200 and 300 μM with PGRP-S resulted in a K_D of 1.89 x 10^{-4} M.

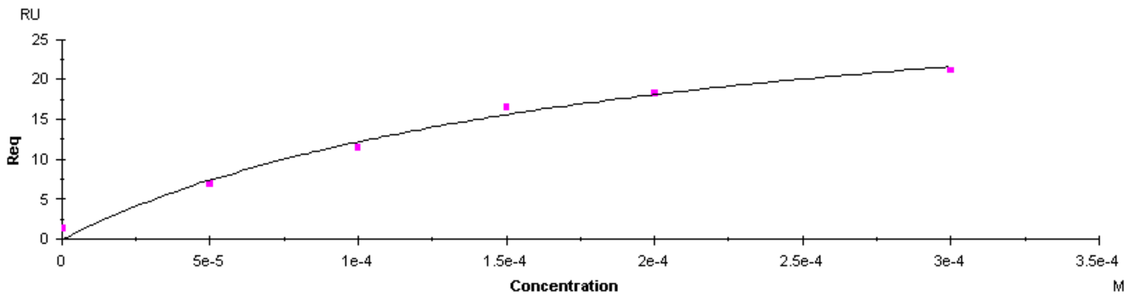


Figure 3.19: Non-linear steady-state affinity analysis for MPP-Lys (**6**) with PGRP-S.

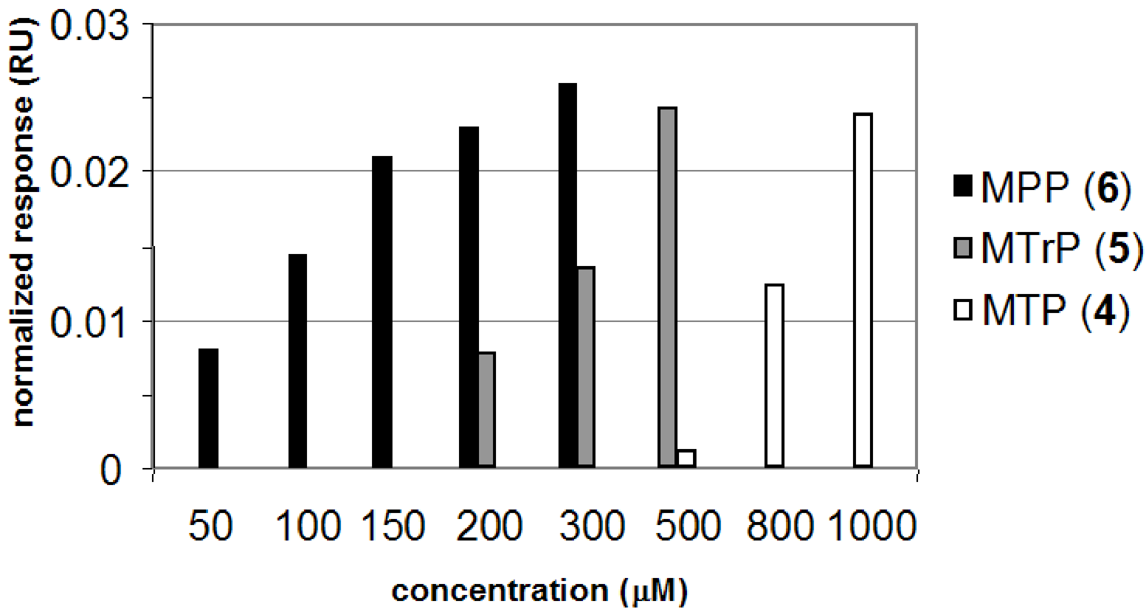


Figure 3.20: Comparative binding analysis of Lysine containing MTP **(4)**, MTrP **(5)** and MPP **(6)** with immobilized PGRP-S (5,000 RU). Concentrations of 50, 100, 150, 200, 300, 500, 800 and 1000 μM of each analyte were passed over the surface for 5 minutes. The average RU from 270 to 300 seconds was divided by respective molecular weight of each analyte to achieve a normalized RU to allow a comparison of affinity through the bar diagram. MDP (**3**) and Lysine-pentapeptide (**1**) had no detectable binding up to 1mM concentrations.

The normalized data for the Dap containing compounds 7-**9** showed a similar trend to that observed of the lysine-containing derivatives **4-6** whereby MTrP-Dap **8** and MPP-Dap **9** were significantly better ligands than MTP-DAP **7**. However, unlike the binding with PGRP-IαC, which displayed similar binding affinities for the lysine- and DAP-type part structures, PGRP-S demonstrated significantly higher affinities for the DAP-

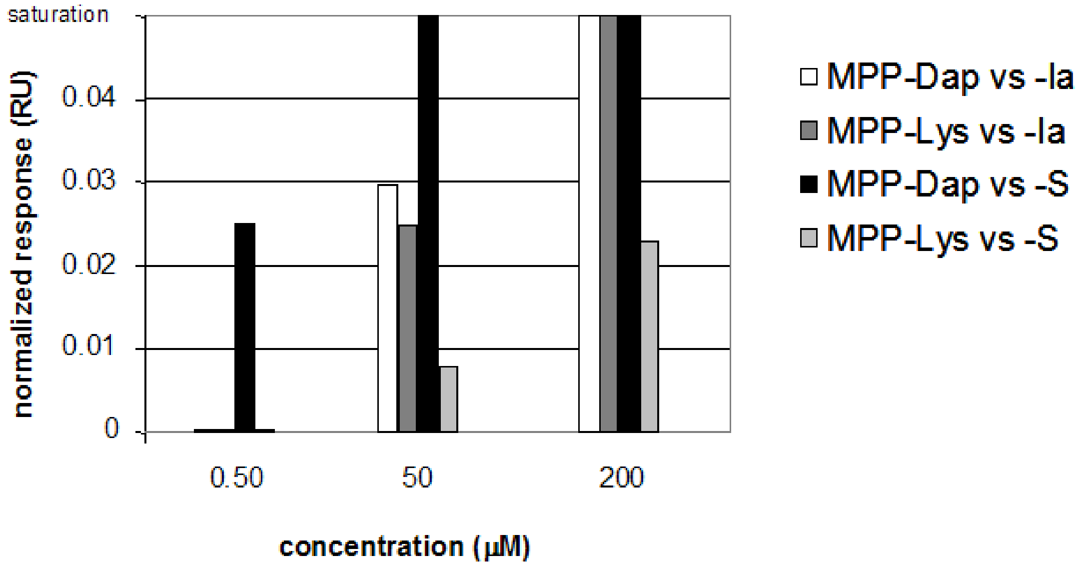


Figure 3.21: Comparative binding analysis of Lysine containing MPP (**6**) and DAP containing MPP (**9**) with PGRP-Iα and PGRP-S at concentrations of 0.50, 50 and 200 μM.

containing fragments. Normalized data for MPP **6** and MPP-DAP **9** with both proteins at concentrations from 0.50 to 200 μM illustrates the highly selective nature of PGRP-S for PGN part structures while PGRP-Iα does not discriminate between the Lysine and DAP

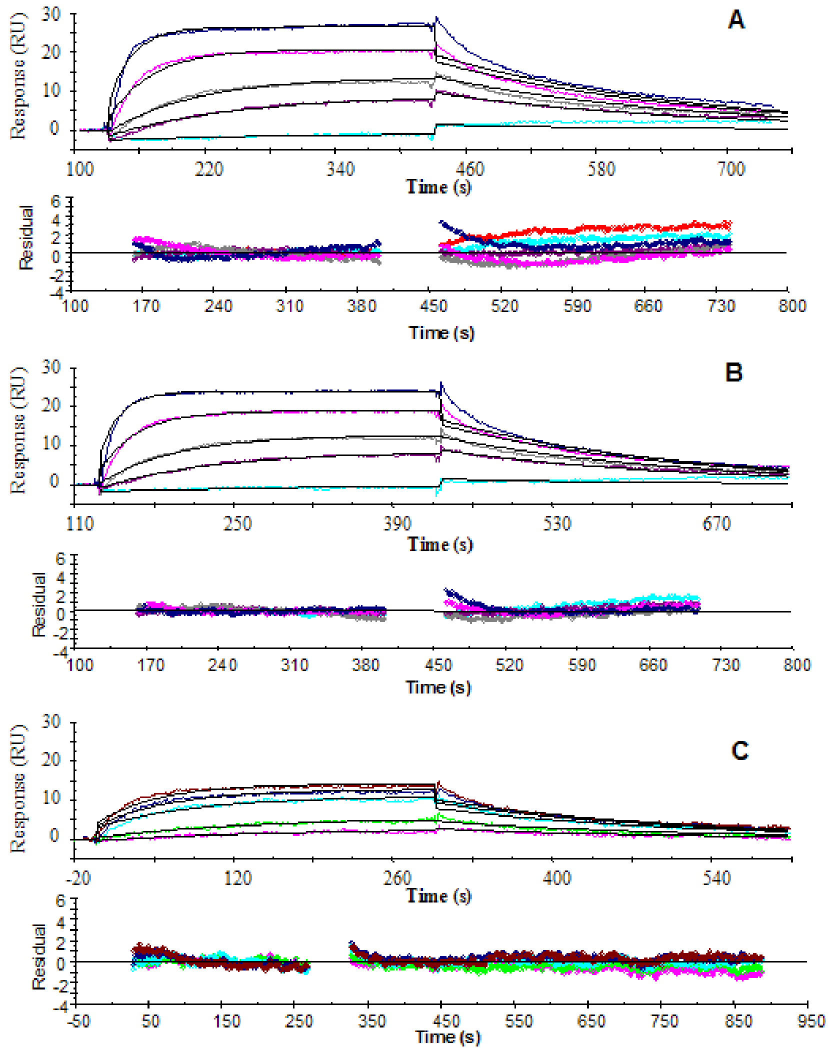


Figure 3.22. Sensorgrams representing the concentration dependent kinetic analysis of DAP containing MPP **(9)**, MTrP **(8)** and MTP **(7)** with immobilized PGRP-S (2,700 RU). Simultaneous kinetic analysis of (**A**) MPP-DAP (**9**) at concentrations of 10, 100, 200, 500 and 1000 nM resulted in k*a*, 4.13 x 10^4 $M^{-1}s^{-1}$; k*d*, 4.29 x 10^{-3} s^{-1}; K_D, 1.04 x 10^{-7} M (**B**) MTrP-DAP (**8**) at concentrations of 10, 100, 200, 500 and 1000 nM resulted in k*a*, 5.17 x 10^4 $M^{-1}s^{-1}$; k*d*, 4.78 x 10^{-3} s^{-1}; K_D of 9.24 x 10^{-8} M and (**C**) MTP-DAP (**7**) at concentrations of 100, 200, 500, 750 and 1000 nM resulted in k*a*, 1.41 x 10^4 $M^{-1}s^{-1}$; k*d*,

$4.61 \times 10^{-3}\ s^{-1}$; K_D of 3.26×10^{-7} M. Corresponding residual values are plotted below the individual sensorgrams.

containing fragments (Figure 3.21). Due to the high affinity for this set of analytes, a lower-density surface containing 2700 RU of PGRP-S was prepared for kinetic analysis. The kinetic wizard method and simultaneous kinetic fitting gave the following dissociation constants: MPP-Dap K_D =104 nM, MTrP-Dap =92 nM, and MTP-Dap =326 nM (Figure 3.22, Table 3.1).

3.5 Discussion

Although the structure of the carbohydrate backbone of peptidoglycan is preserved among all bacteria, considerable structural variability exists in the peptide moiety (216). The archetypal stem peptide of gram-positive bacteria is L-alanine-γ-D-glutamate-L-lysine-D-Ala. In the case of Gram-negative bacteria and Gram-positive *bacilli*, *m*-diaminopimelic acid (DAP-type) is normally found as the third amino acid. *Drosophila* PGRP-SA has been shown to interact with lysine-type PGN activating the Toll receptor pathway (217), whereas PGRP-LC and PGRP-LE recognize Dap-type PGN activating the *Imd*/*Relish* pathway (218-222). Previous studies have shown that each of the four mammalian PGRPs is able to bind peptidoglycan. However, possible selectivities for lysine- or Dap-type PGN have either not been determined or remain controversial.

In this study, the complexation of human PGRP-IαC and PGRP-S with a range of synthetic PGN fragments, containing either lysine- or DAP as the third amino acid, have been studied by SPR. A dissociation constant of K_D = 62 μM for MPP-Lys **6** was determined by a fitting of steady state conditions. Furthermore, the normalized data of the lysine-containing compounds **4-6** indicate that MTrP **5** and MPP **6** have similar affinities whereas a much lower affinity for MTP **4** was measured. In addition, peptides **1** (Lys-PP) and **2** (DAP-PP) and MDP (**3**) exhibited no binding at 1mM concentration.

Recently, the crystal structure of PGRP-IαC ligated with MTP **4** was reported at 2.3Å (Figure 3.23) (235). In this complex, the tripeptide stem of MTP was held in an extended conformation at the deep end of the binding groove, whereas the MurNAc moiety lies in a pocket in the middle of the groove, with the pyranose ring oriented perpendicular to the base of the pocket. The structure indicates that the protein can accommodate a fourth D-Ala residue making contacts with Gln-261, Tyr-266 and Asn-269. The SPR data reported here demonstrate that these proposed interactions contribute significantly to binding. By contrast, D-Ala at position 5 is expected to extend beyond the binding groove. Therefore this amino acid should contribute little to the binding, an observation consistent with the SPR data. The observation that the pentapeptide **1** does not bind with PGRPIαC indicates that the interactions with the muramic acid moiety are critical for binding. In this respect, the lactyl moiety forms a hydrogen bond with Y242 whereas the acetamido of the

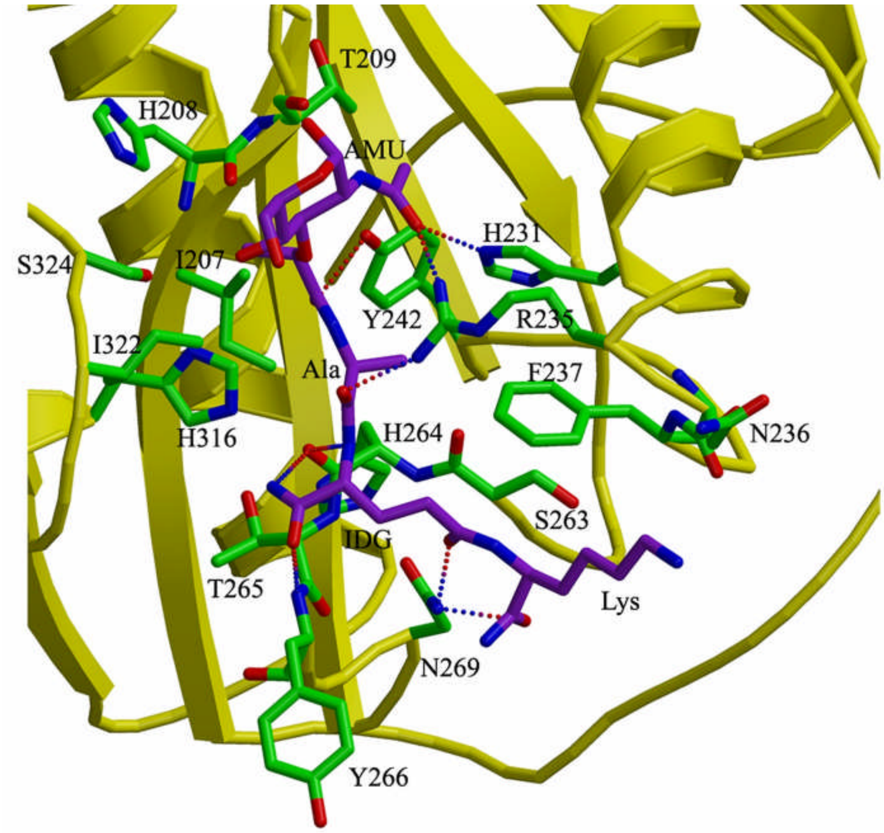


Figure 3.23: Intermolecular contacts in the PGRP-IαC-MTP-Lysine complex. Stereoview of interactions between PGRP-IαC and MTP-Lysine (**4**) at the PGN-binding site. MTP-Lysine is shown in purple, PGRP-IαC in yellow and contacting residues in green. Hydrogen bonds are shown as dashed lines; residues forming van der Waals contacts with MTP-Lysine are also highlighted.

saccharide moiety interacts with His2231 and R235. SPR experiments with Dap-containing fragments **7-9** revealed that they complex with similar affinities as compared to the Lys-containing compounds **4-6**.

The complexation of PGRP-S with Dap-containing fragments **7-9** exhibited high affinities with dissociation constants in the nanomolar range. The affinity trend was similar to that observed for the binding of PGRP-IαC with Dap-containing compounds **7-9**. Thus, MTP-Dap (**7**) was the smallest fragment to be recognized and significantly higher affinities were observed for MTrP-Dap (**8**) and MPP-Dap (**9**). Interestingly, much lower affinities were measured for the interactions of PGRP-S with the lysine-type compounds **4**-**6**. Thus, while only a minimal selectivity was observed for binding of PGRP-IαC with lysine- and DAP-containing compounds, PGRP-S displayed a significant preference for DAP-containing PGN fragments.

The PGRP-IαC – MTP complex shows that Asn-236 and Phe-237 form a number of van der Waals contacts with the side chain of L-lysine. Sequence alteration at these two positions may account for the discriminatory ability of different PGRPs towards Lys type and Dap type PGNs. For example, the corresponding sequence in *Drosophila* PGRP-LCx and PGRP-LE, which recognizes Dap type PGNs (152,194), is Gly–Trp. PGRP-S also has Gly^{89} and Trp^{90} at these positions.

The structural difference between lysine and Dap is an extra carboxylic acid at the εC of the side chain. This chemical difference does not allow the different PGRPs to exhibit absolute discriminatory ability towards one type of PGN. For example, a certain degree of cross-reactivity is also exhibited by *Drosophila* PGRP-SA and mouse PGRP-L, which hydrolyses PGNs from both Gram-positive and negative sources (227). Similarly, in the case of PGRP-IαC, despite having Asn-236 and Phe-237, the protein can recognize both Lys and Dap-type PGNs with similar affinities.

Selectivity for lysine or Dap-containing peptidoglycan fragments has been observed for other pattern recognition receptors. For example, transfection studies have shown that NOD2 recognized MDP and muramyltripeptide containing lysine as the third amino acid (195). On the other hand, NOD1 senses DAP containing peptidoglycan (238-239). The structure activity relationships determined for the NOD proteins shows a different profile from what has been determined for PGRP-Iα and PGRP-S. In this respect, the NOD proteins do not respond to muramyl tetra- and muramyl pentapeptide structures. It appears that the different pattern recognition receptors for peptidoglycan have evolved in such a way that they can recognize different part structures.

References

[1] A. Varki, *Glycobiology* **1993**, *3*, 97.

[2] J. Dennis, M. Granovsky and C. Warren, *Biochim. Biophys. Acta* **1999**, *1473*, 21.

[3] S. Hakomori, *Cancer Res.* **1985**, *45*, 2405.

[4] J. Kellerman, F. Lottspeich, A. Henschen and W. Muller-Esterl, *Eur. J. Biochem.* **1986**, *154*, 471.

[5] J. Montreuil, *Adv. Carbohydr. Chem. Biochem.* **1980**, *37*, 157.

[6] N. Sharon, *Trends Biochem. Sci.* **1984**, *9*, 198.

[7] S. Hakomori and A. Kombata, *The Antigens*; *Sela, M., Ed*., **1974**; *2,* 79.

[8] L. Lasky, *Science*, **1992**, *258*, 964.

[9] F. Ofosu, I. Danishefsky and J. Hirsh, Eds. *Heparin and Related Polysaccharides. New York: NY Acad. Sci.* **1989**, 501.

[10] J. Dennis, *Semin. Cancer Biol.* **1991**, *2,* 411.

[11] J. Dennis, S. Laferte, S. Yagel, and M. Breitman, *Cancer Cells* **1989**, *1*, 87.

[12] J. Baenziger, *FASEB J*. **1994**, *8*, 1019.

[13] S. Kornfeld, *FASEB J.* **1987**, *1,* 462.

[14] Y. Lee and R. Lee, *Acc. Chem. Res.* **1995**, *8*, 321.

[15] R. Laine, *Pure Appl. Chem.* **1997**, *69*, 1867.

[16] F. Quiocho, *Annu. Rev. Biochem.* **1986**, *55*, 287.

[17] W. Weis and K. Drickamer, *Annu. Rev. Biochem.* **1996**, *65*, 441.

[18] H. Lis and N. Sharon, *Chem. Rev.* **1998**, *98*, 637.

[19] E. Toone, *Curr. Opin. Struct. Biol.* **1994**, *4*, 719.

[20] R. Loris, P. Stas and L. Wyns, *J. Biol. Chem.* **1994**, *269*, 26722.

[21] K. Drickamer and M. Taylor, *Annu. Rev. Cell. Biol.* **1993**, *9*, 237.

[22] L. Tremblay and A. Herscovics, *Glycobiology* **1999**, *9*, 1073.

[23] K. Ng, S. Park-Snyder and W. Weis, *Biochem.* **1998**, *37*, 17965.

[24] S. Faham, R. Lindhardt and D. Rees, *Curr. Opin. Struct. Biol.* **1998**, *8*, 578.

[25] T. Corfield, *Glycobiology* **1992**, *2*, 509.

[26] F. Quiocho, *Pure Appl. Chem.* **1989**, *61*, 1293.

[27] F. Quiocho, J. Spurlino and L. Rodseth, *Nat. Struct. Biol.* **1997**, *5*, 997.

[28] C. Malet and O. Hindsgaul, *J. Org. Chem.* **1996**, *61*, 4649.

[29] R. Roy, *Curr. Opin. Struct. Biol.* **1996**, *6*, 692.

[30] Y. Lee, *FASEB J.* **1992**, *6*, 3193.

[31] P. Kitov and D. Bundle, *J. Am. Chem. Soc.* **2003**, *125*, 16271.

[32] S. Shenoy, L. Barrientos, D. Ratner, B. O'Keefe, P. Seeberger, A. Gronenborn and M. Boyd, *Chem. Biol.* **2002**, *9*, 1109.

[33] T. Dam, H. Gabius, S. Andre, H. Kaltner, M. Lensch and C. Brewer, *Biochem.* **2005**, *44*, 12564.

[34] R. Lee and Y. Lee, *In Y. Lee and R. Lee Eds. Neoglycoconjugates: Preparation and Applications, San Diego, CA, Academic Press* **1994**, 23.

[35] M. Page and W. Jencks, *Proc. Nat. Acad. Sci. USA* **1971**, *68*, 1678.

[36] R. Liang, J. Loebach, N. Horan, M. Ge, C. Thompson, L. Yan and D. Kahne *Proc. Nat. Acad. Sci. USA* **1997**, *94*, 10554.

[37] Y. Shinohara, Y. Hasegawa, H. Kaku and N. Shibuya, *Glycobiology* **1997**, *7*, 1201.

[38] J. Gestwicki, L. Strong and L. Kiessling, *Chem. Biol.* **2000**, *7*, 583.

[39] K. Boehme and T. Compton, *J. Virol.* **2004**, *78*, 7867.

[40] J. Rini, *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 551.

[41] D. Mandal, N. Kishore and C. Brewer, *Biochem.* **1994**, *33*, 1149.

[42] D. Mandal, L. Bhattacharyya, S. Koenig, R. Brown III, S. Oscarson and C. Brewer, *Biochem.* **1994**, *33*, 1157.

[43] M. Mammen, S. Choi and G. Whitesides, *Angew. Chem. Int. Ed.* **1998**, *37*, 2755.

[44] S. Borman, *Chem. Eng. News* **2000**, *78*, 48.

[45] D. Mann and L. Kiessling, *In P. Wang and C. Bertozzi, Eds. Glycochemistry; Principles, Synthesis and Applications, New York: Marcel Dekker, Inc.* **2001**, 221.

[46] L. Kiessling and N. Pohl, *Chem. Biol.* **1996**, *3*, 71.

[47] Y. Lee, R. Townsend, M. Hardy, J. Lonngren, J. Arnarp, M. Haraldsson and H. Lonn, *J. Biol. Chem.* **1983**, *258*, 199.

[48] A. Spaltenstein and G. Whitesides, *J. Am. Chem. Soc.* **1991**, *113*, 686.

[49] M. Matrosovich, L. Mochalova, V. Marinina, N. Byramova and N. Bovin, *FEBS Lett.* **1990**, *272*, 209.

[50] R. Roy, M. Baek and K. Rittenhouse-Olson, *J. Am. Chem. Soc.* **2001**, *123*, 1809.

[51] D. Revell, J. Knight, D. Blyth, A. Haines and D. Russell, *Langmuir* **1998**, *14*, 4517.

[52] G. Newkome, Z. Yao, G. Baker and V. Gupta, *J. Org. Chem.* **1985**, *50*, 2003.

[53] D. Tomalia, H. Baker, J. Dewald, M. Hall, G. Kallos, S. Martin, J. Roeck, J. Ryder and P. Smith, *Polymer J.* **1985**, *17*, 117.

[54] K. Tsutsumiuchi, K. Aoi and M. Okada, *Polymer J.* **1999**, *31*, 935.

[55] R. Schauer, *Adv. Carbohydr. Chem. Biochem.* **1982**, *40*, 131.

[56] W. Reutter, E. Kottgen, C. Bauer and W. Gerok, *Cell. Biol. Monogr.* **1982**, *10*, 263.

[57] R. Schauer Ed., *Sialic Acids: Chemistry, Metabolism and Function, Wien, NY. Springer-Verlag* **1982**, *10*.

[58] A. Rosenberg Ed., *Biology of Sialic Acids, New York, Plenum Press* **1995**.

[59] C. Traving and R. Schauer, *Cell. Mol. Life Sci.* **1998**, *54*, 1330.

[60] A. Corfield, J. Michalski and R. Schauer, *In Sialidases and Sialidosis: Perspectives in Inherited Metabolic Diseases, G. Tettamanti, P. Durand and S. DiDonato Eds.,* **1981**, 3.

[61] A. Savage, J. Donohue, C. Koeleman and D. van den Eijnden, *Eur. J. Biochem.* **1990**, *193*, 837.

[62] M. Engstler, R. Schauer and R. Brun, *Acta Trop.* **1995**, *59*, 117.

[63] S. Schenkman, D. Eichinger, M. Periera and V. Nussenzweig, *Annu. Rev. Microbiol.* **1994**, *48*, 499.

[64] E. Saavedra, M. Herrera, W. Gao, H. Uemera and M. Periera, *J. Exp. Med.* **1999**, *190*, 1825.

[65] G. Boons and A. Demchenko, *Chem. Rev.* **2000**, *100*, 4539.

[66] S. Kelm and R. Schauer, *Int. Rev. Cytol.* **1997**, *175*, 137.

[67] R. Schauer, A. Shukla, C. Schroder and E. Muller, *Pure Appl. Chem.* **1984**, *56*, 907.

[68] J. Morgenthaler; W. Kemmner and R. Brossmer, *Biochem. Biophys. Res. Commun.* **1990**, *171*, 860-866.

[69] R. Schauer, *Methods in Enzymol.* **1987**, *138*, 132.

[70] H. Muller, *Dtsch. Med. Wschr.* **1974**, *99*, 1933.

[71] G. Born and W. Palinski, *J. Physiol.* **1989**, *419*, 169.

[72] R. Schauer, *Nachrichten aus Chemie, Technik und Laboratorium* **1992**, *40*, 1227.

[73] J. Montreil, *In Comprehensive Biochemistry, E. Florkin Ed., Amsterdam, Elsevier* **1982**, *19B*.

[74] V. Bhavanandan, N. Ringler and D. Channe Gowda, *Glycobiology* **1998**, *8*, 1077.

[75] G. Taylor, *Curr. Opin. Struct. Biol.* **1996**, *6*, 830.

[76] G. Davies, M. Sinnott, and S. Withers, *In M. Sinnott Ed. Comprehensive Biological Catalysis*, *San Diego: Academic Press*, **1998**, 119.

[77] J. Watson, S. Newstead, V. Dookhun, G. Taylor and A. Bennet, *FEBS* **2004**, *577*, 265.

[78] D. Koshland Jr., *Biol. Rev.* **1953**, *28*, 416.

[79] A. Vasella, G. Davies and M. Bohm, *Curr. Opin. Struct. Biol.* **2002**, *6*, 619.

[80] J. Varghese, W. Laver and P. Colman, *Nature (London)* **1983**, *303*, 35.

[81] S. Crennell, E. Garman, W. Laver, E. Vimr and G. Taylor, *Proc. Natl. Acad. Sci.* **1993**, *90*, 9852.

[82] S. Crennell, T. Takimoto, A. Portner and G. Taylor, *Nat. Struct. Biol.* **2000**, *7*, 1068.

[83] A. Buschiazzo, M. Amaya, M. Cremona, A. Frasch and P. Alzari, *Mol. Cell* **2002**, *10*, 757.

[84] D. Bratosin, J. Mazurier, H. Debray, M. Lecocq, B. Boily, C. Alonso, M. Moisei, C. Montas and J. Montreuil, *Glyconj. J.* **1995**, *12*, 258.

[85] M. Peter, S. Hellbardt, R. Schwartzalbiez, M. Westendorp, H. Walczak, G. Moldenhauer, M. Grell and P. Krammer, *Cell Death Differ.* **1995**, *2*, 163.

[86] Y. Pilatte, J. Bignon and C. Lambre, *Glycobiology* **1993**, *3*, 201.

[87] M. Saito and R. Yu, *J. Neurosci. Res.* **1993**, *36*, 127.

[88] G. Hirst, *Science* **1941**, *94*, 22.

[89] P. Palese, K. Tobita, M. Ueda and R. Compans, *Virol.* **1974**, *61*, 397.

[90] L. Hoskins, E. Boulding, T. Gerken, V. Harouny and M. Kriaris, *Microb. Ecol. Health Dis.* **1992**, *5*, 193.

[91] H. Muller, *Infektionen. Zbl. Bakt. Hyg.* **1976**, *235*, 106.

[92] J. Galen, J. Ketley, A. Fasano, S. Richardson, S. Wasserman and J. Kaper, *Infect. Immunol.* **1992**, *60,* 406.

[93] A. Corfield, S. Wagner, J. Clamp, M. Kriaris and L. Hoskins, *Infect. Immunol.* **1992**, *60*, 3971.

[94] M. Teufel, P. Roggentin and R. Schauer, *Biol. Chem. Hoppe-Seyler* **1989**, *370*, 435.

[95] A. Fraser and R. Brown, *J. Med. Microbiol.* **1981**, *14*, 63.

[96] C. Guzman, M. Plate and C. Pruzzo, *FEMS Microbiol. Lett.* **1990**, *59*, 187.

[97] A. Brown, *J. Chem. Soc.* **1902**, *81*, 373.

[98] G. Briggs and J. Haldane, *Biochem. J.* **1925**, *19*, 383.

[99] V. Henri, *Lois Generales de l'action des diastases, Hermann, Paris* **1903.**

[100] L. Michaelis and M. Menten, *Biochem. Z.* **1913**, *49*, 333.

[101] J. Hoffman, F. Kafatos, C. Janeway and R. Ezekowitz, *Science* **1999**, *284*, 1313.

[102] A. Aderem and R. Ulevitch, *Nature* **2000**, *406*, 782.

[103] C. Janeway and R. Medzhitov, *Annu. Rev. Immunol.* **2002**, *20*, 197.

[104] D. Philpott and S. Girardin, *Mol. Immunol.* **2004**, *41*, 1099.

[105] G. Barton and R. Medzhitov, *Science* **2003**, *300*, 1524.

[106] K. Anderson, L. Bokla and C. Nusslein-Volhard, *Cell* **1985**, *42*, 791.

[107] B. Lemaitre, E. Nicholas, L. Michaut, J. Reichhart and J. Hoffmann, *Cell* **1996**, *86*, 973.

[108] B. Lemaitre, J. Reichhart and J. Hoffmann, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 14614.

[109] F. Rock, G. Hardiman, J. Timans, R. Kastelein and J. Bazan, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 588.

[110] A. Poltorak, X. He, I. Smirnova, M. Liu, C. Van Huffel, X. Du, D. Birdwell, E. Alejos, M. Silva, C. Galanos, M. Freudenberg, P. Ricciardi-Castagnoli, B. Layton and B. Beutler, *Science* **1998**, *282*, 2085.

[111] K. Takeda, T. Kaisho and S. Akira, *Annu. Rev. Immunol.* **2003**, *21*, 335.

[112] R. Medzhitov, P. Preston-Hurlburt and C. Janeway, *Nature* **1997**, *388*, 394.

[113] S. Akira, *Curr. Opin. Immunol.* **2003**, *15*, 5.

[114] L. O'Neill, K. Fitzgerald and A. Bowie, *Trends Immunol.* **2003**, *24*, 286.

[115] M. Tsan, *Seminars Cancer Biol.* **2006**, *16*, 32.

[116] S. Girardin, I. Boneca, L. Carneiro, A. Antignac, M. Jehanno, J. Viala, K. Tedin, M. Taha, A. Labigne, U. Zahringer, A. Coyle, P. DiStefano, J. Bertin, P. Sansonetti and D. Philpott, *Science* **2003**, *300*, 1584.

[117] S. Girardin, I. Boneca, J. Viala, M. Chamaillard, A. Labigne, G. Thomas, D. Philpott and P. Sansonetti, *J. Biol. Chem.* **2003**, *278*, 8869.

[118] N. Inohara, Y. Ogura, A. Fontalba, O. Gutierrez, F. Pons, J. Crespo, K. Fukase, S. Inamura, S. Kusmoto, M. Hashimoto, S. Foster, A. Moran, J. Fernandez-Luna and G. Nunez, *J. Biol. Chem.* **2003**, *278*, 5509.

[119] S Girardin, R. Tournebize, M. Mavris, A. Page, X. Li, G. Stark, J. Bertin, P. DiStefano, M. Yaniv, P. Sansonetti and D. Philpott, *EMBO Rep.* **2001**, *2*, 736.

[120] H. Yoshida, K. Kinoshita and M. Ashida, *J. Biol. Chem.* **1996**, *271*, 13854.

[121] R. Dziarski, *Mol. Immunol.* **2004**, *40*, 877.

[122] A. Takehana, T. Katsuyama, T. Yano, Y. Oshima, H. Takeda, T. Aigaki and S. Kurata, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13705.

[123] T. Michel, J. Reichhart, J. Hoffmann and J. Royet, *Nature* **2001**, *414*,756.

[124] K. Choe, T. Werner, S. Stoven, D. Hultmark and K. Anderson, *Science* **2002**, *296*, 359.

[125] M. Gottar, V. Gobert, T. Michel, M. Belvin, G. Duyk and J. Hoffmann, *Nature* **2002**, *416*, 640.

[126] M. Ramet, P. Manfruelli, A. Pearson, B. Mathey-Prevot and R. Ezekowitz, *Nature* **2002**, *416*, 644.

[127] D. Kang, G. Liu, A. Lundstrom, E. Gelius and H. Steiner, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 10078.

[128] C. Liu, Z. Xu, D. Gupta and R. Dziarski, *J. Biol. Chem.* **2001**, *276*, 34686.

[129] S. Wright, *J. Immunol.* **1995**, *155*, 6.

[130] R. Delude, R. Savedra, H. Zhao, R. Thieringer, S. Yamamoto, M. Fenton and D. Golenbock, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9288.

[131] S. Wright, R. Ramos, P. Tobias, R. Ulevitch and J. Mathison, *Science* **1990**, *249*, 1431.

[132] A. Haziot, E. Ferrero, F. Kontgen, N. Hijiya, S. Yamamoto, J. Silver, C. Stewart and S. Goyert, *Immunity* **1996**, *4*, 407.

[133] R. Dziarski, R. Tapping and P. Tobias, *J. Biol. Chem.* **1998**, *273*, 8680.

[134] T. Espevik, M. Otterlei, G. Skjak-Braek, L. Ryan, S. Wright and A. Sundan, *Eur. J. Immunol.* **1993**, *23*, 255.

[135] M. Soell, E. Lett, F. Holveck, M. Scholler, D. Wachsmann and J. Klein, *J. Immunol.* **1995**, *154*, 851.

[136] M. Cleveland, J. Gorham, T. Murphy, E. Toumanen and K. Murphy, *Infect. Immunol.* **1996**, *64*, 1906.

[137] R. Dziarksi, *Cell. Mol. Life Sci.* **2003**, *60*, 1793.

[138] J. Pugin, C. Schurer-Maly, D. Leturq, A. Moriarty, R. Ulevitch and P. Tobias, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 2744.

[139] R. Dziarski, S. Viriyakosol, T. Kirkland and D. Gupta, *Infect. Immunol.* **2000**, *68*, 5254.

[140] K. Aslan, J. Lakowicz and C. Geddes, *Curr. Opinion Chem. Biol.* **2005**, *9*, 538-544.

[141] T. Turbadar, *Proc. Phys. Soc.* **1959**, *73*, 40.

[142] A. Otto, *Z. Phys.* **1968**, *216*, 398.

[143] *Biacore Basics.* **2004**, *1,* 1-28.

[144] I. Langmuir, *J. Am. Chem. Soc.* **1918**, *40*, 1361.

[145] G. Scatchard, *Am. N.Y. Acad. Sci*. **1949**, *51*, 660.

[146] R. Karlsson, A. Michaelsson and L. Mattsson, *J. Immunol. Methods* **1991**, *145,* 229.

[147] D. O'Shannessy, M. Brigham-Burke, K. Soneson, P. Hensley and I. Brooks, *Anal. Biochem.* **1993**, *212*, 457.

[148] R. Glaser, *Anal. Biochem.* **1993**, *212*, 152.

[149] D. O'Shannessy and D. Wiznor, *Anal. Biochem.* **1996**, *236*, 275.

[150] E. Duverger, N. Frison, A. Roche and M. Monsigny, *Biochimie*, **2003**, *85*, 167.

[151] S. Kumar, A. Roychowdhury, B. Ember, Q. Wang, R. Guan, R. Mariuzza and G. Boons, *J.Biol. Chem.* **2005**, *280*: 37005.

[152] T. Corfield, *Glycobiology* **1992**, *2*, 509.

[153] J. Galen, J. Ketle, A. Fasano, S. Richardson, S. Wasserman and J. Kaper, *Infect. Immun.* **1992**, *60,* 406.

[154] D. Schneider and C. Parker, *J. Infect. Dis.* **1982**, *145,* 474.

[155] R. Schauer, *Trends Biochem. Sci.* **1985**, *10*, 357.

[156] Y. Pilatte, J. Bignon and C. Lambre, *Glycobiology* **1993**, *3*, 201.

[157] M. Cantz and B. Ulrich-Bott, *J. Inherit. Metab. Dis.* **1990**, *13*, 523.

[158] A. Gaskell, S. Crennell and G. Taylor, *Structure* **1995**, *3*, 1197.

[159] C. Khosla and P. Harbury, *Nature* **2001**, *409*, 247.

[160] G. Black, G. Hazlewood, S. Millward-Sadler, J. Laurie and H. Gilbert, *Biochem. J.* **1995**, *307*, 191.

[161] S. Thobhani, B. Ember, A. Siriwardena and G. Boons, *J. Am. Chem. Soc.* **2002**, *125*, 7154.

[162] Y. Lee and R. Lee, *Acc. Chem. Res.* **1995**, *28*, 321.

[163] L. Kiessling, J. Gestwick and L. Strong, *Curr. Opin. Chem. Biol.* **2000**, *4*, 696.

[164] J. Lundquist and E. Toone, *Chem. Rev.* **2002**, *102,* 555.

[165] R. Liang, J. Loebach, N. Horan, M. Ge, C. Thompson, L. Yan and D. Kahne, *Proc. Natl. Acad. Sci. USA* **1997**, *97*, 10554.

[166] J. Klemm, S. Schreiber and G. Crabtree, *Annu. Rev. Immunol.* **1998**, *16*, 569.

[167] A. Siriwardena, M. Jorgensen, M. Wolfert, M. Vandenplas, J. Moore and G. Boons, *J. Am. Chem. Soc.* **2001**, *123*, 8145.

[168] C. Traving and R. Schauer, *Cell. Mol. Life Sci.* **1998**, *54*, 1330.

[169] S. Kelm and R. Schauer, *Int. Rev. Cytol.* **1997**, *175*, 137.

[170] G. Taylor, *Curr. Opin. Struct. Biol.* **1996**, *6*, 830.

[171] L. Warren, *J. Biol. Chem.* **1959**, *234*, 1971.

[172] J. Tallman, P. Fishman and R. Henneberry, *Arch. Biochem. Biophys.* **1977**, *182*, 556.

[173] J. Goddard and J. Reymond, *Trends Biotechnol.* **2004**, *22*, 363.

[174] M. Potier, L. Mameli, M. Belisle, L. Dallaire and S. Melancon, *Anal. Biochem.* **1979**, *94*, 287.

[175] Nakamura, M., Hara, S., Yamaguchi, M., Takemori, Y., and Ohkura, Y. *Chem. Pharm. Bull.* **1987**, *35*, 687.

[176] S. Crennel, E. Garman, G. Laver, E. Vimr and G. Taylor, *Structure* **1994**, *2*, 535.

[177] S. Crennel, E. Garman, C. Philippon, A. Vasella, W. Laver, E. Vimr and G. Taylor, *J. Mol. Biol.* **1996**, *259*, 264.

[178] D. Schmidt, B. Sauerbrei and J. Thiem, *J. Org. Chem.* **2000**, *65*, 8518.

[179] M. Klefel and M. von Itzstein, *Chem. Rev.* **2002**, *102*, 471.

[180] S. Newstead, J. Watson, A. Bennet and G. Taylor *Biol. Crystal.* **2005**, *61*, 1483.

[181] I. Moustafa, H. Connaris, M. Taylor, V. Zaitsev, J. Wilson, M. Kiefel, M. von Itzstein and G. Taylor, *J. Biol. Chem.* **2004**, *279*, 40819.

[182] A. Boraston, B. McLean, J. Kormos, M. Alam, N. Gilkes, C. Haynes, P. Tomme, D. Kilburn and R. Warren, *In Gilbert, H., Davies, G., Henrissat, B., and Svensson, B. eds. Recent Advances in Carbohydrate Bioengineering*, Cambridge: Royal Society of Cambridge. **1999**, 202.

[183] W. Turnbull and R. Field, *J. Chem. Soc. Perkin Trans. 1* **2000**, 1859.

[184] R. Lemieux, A. Pavia, J. Martin and K. Watanabe, *Can. J. Chem.* **1969**, *47*, 4427.

[185] A. Boraston, D. Bolam, H. Gilbert and G. Davies, *Biochem. J.* **2004**, *382*, 769.

[186] M. Hardy and R. Townsend, *Methods Enzymol.* **1994**, *230*, 208.

[187] K. Kartha, A. Kameyama, M. Kiso and Y. Hasegawa, *J. Carbohydr. Chem.* **1989,** *8,* 145.

[188] T. Murase, K. Kartha, M. Kiso and Y. Hasegawa, *Carbohydr. Res.* **1989,** *195,* 134.

[189] N. Kochetkov, N. Byramova, Y. Tsvetkov and L. Backinowsky, *Tetrahedron* **1985**, *41*, 3363.

[190] R. Lemieux and R. Stick, *Aust. J. Chem.* **1975**, *28*, 1799.

[191] A. Hasegawa, J. Nakamura and M. Kiso, *J. Carbohydrate Chem.* **1986**, *5*, 11.

[192] H. Lonn, *Carbohydrate Res.* **1985**, *139*, 105.

[193] M. Mammen, G. Dahmann and G. Whitesides, *J. Med. Chem.* **1995**, *38*, 4179.

[194] R. Binkley, M. Ambrose and D. Hehemann, *J. Org. Chem.* **1980**, *45*, 4387.

[195] J. Hoffmann, *Nature* **2003**, *426*, 33.

[196] C. Janeway and R. Medzhitov, *Annu. Rev. Immunol.* **2002**, *20*, 197.

[197] R. Medzhitov and C. Janeway, *Science* **2002**, *296*, 298.

[198] B. Beutler, *Mol. Immunol.* **2004**, *40*, 845.

[199] E. Van Amersfoort, T. Van Berkel and J. Kuiper, *Clin. Microbiol. Rev.* **2003**, *16*, 379.

[200] B. Beutler, K. Hoebe, X. Du and R. Ulevitch, *J. Leukocyte Biol.* **2003**, *74*, 479.

[201] E. Lien and R. Ingalls, *Crit. Care Med.* **2002**, *30*, S1.

[202] L. O'Neill, *Science* **2004**, *303*, 1481.

[203] W. Check, *Asm News* **2004**, *70*, 317.

[204] L. O'Neill, *Trends Immunol.* **2004**, *25*, 687.

[205] F. Schmitz, J. Mages, A. Heit, R. Lang and H. Wagner, *Eur. J. Immunol.* **2004**, *34*, 2863.

[206] C. Pasare and R. Medzhitov, *Curr. Opin. Immunol.* **2003**, *15*, 677.

[207] M. Tsan and B. Gao, *J. Leukocyte Biol.* **2004**, *76*, 514.

[208] S. Sarkar, K. Peters, C. Elco, S. Sakamoto, S. Pal and G. Sen, *Nat. Struct. Mol. Biol.* **2004**, *11*, 1060.

[209] L. Travassos, S. Girardin, D. Philpott, D. Blanot, M. Nahori, C. Werts and I. Boneca, *EMBO Rep.* **2004**, *5*, 1000.

[210] N. Inohara and G. Nunez, *Nat. Rev. Immunol.* **2003**, *3*, 371.

[211] M. Chamaillard, S. Girardin, J. Viala and D. Philpott, *Cell. Microbiol.* **2003**, *5*, 581.

[212] R. Dziarski, *Mol. Immunol.* **2004**, *40*, 877.

[213] C. Liu, Z. J. Xu, D. Gupta and R. Dziarski, *J. Biol. Chem.* **2001**, *276*, 34686.

[214] D. Kang, G. Liu, A. Lundstrom, E. Gelius and H. Steiner, *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 10078.

[215] T. Werner, G. Liu, D. Kang, S. Ekengren, H. Steiner and D. Hultmark, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13772.

[216] K. Schleifer and O. Kandler, *Bacteriol. Rev.* **1972**, *36*, 407.

[217] T. Michel, J. Reichhart, J. Hoffmann and J. Royet, *Nature* **2001**, *414*, 756.

[218] T. Werner, K. Borge-Renberg, P. Mellroth, H. Steiner and D. Hultmark, *J. Biol. Chem.* **2003**, *278*, 26319.

[219] M. Ramet, P. Manfruelli, A. Pearson, B. Mathey-Prevot and R. Ezekowitz, *Nature* **2002**, *416*, 644.

[220] K. Choe, T. Werner, S. Stoven, D. Hultmark and K. Anderson, *Science* **2002**, *296*, 359.

[221] M. Gottar, V. Gobert, T. Michel, M. Belvin, G. Duyk, J. Hoffmann and D. Ferrandon, J. Royet, *Nature* **2002**, *416*, 640.

[222] F. Leulier, C. Parquet, S. Pili-Floury, J. Ryu, M. Caroff, W. Lee, D. Mengin-Lecreulx and B. Lemaitre, *Nat. Immunol.* **2003**, *4*, 478.

[223] X. Cheng, X. Zhang, J. Pflugrath and F. Studier, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4034.

[224] P. Mellroth, J. Karlsson and H. Steiner, *J. Biol. Chem.* **2003**, *278*, 7059.

[225] Z. Wang, X. Li, R. Cocklin, M. Wang, M. Wang, K. Fukase, S. Inamura, S. Kusumoto, D. Gupta and R. Dziarski, *J. Biol. Chem.* **2003**, *278*, 49044.

[226] E. Gelius, C. Persson, J. Karlsson and H. Steiner, *Biochem. Biophys. Res. Commun.* **2003**, *306*, 988.

[227] R. Dziarski, K.Platt, E. Gelius, H. Steiner and D. Gupta, *Blood* **2003**, *102*, 689.

[228] C. Tydell, N. Yount, D. Tran, J. Yuan and M. Selsted, *J. Biol. Chem.* **2002**, *277*, 19658.

[229] R. Guan, E. Malchiodi, Q. Wang, P. Schuck and R. Mariuzza, *J. Biol. Chem.* **2004**, *279*, 31873.

[230] P. Sieber, *Tetrahedron Lett.* **1987**, *28*, 2107.

[231] A. Chowdhury, A. Siriwardena and G. Boons, *Tetrahedron Lett.* **2002**, *43*, 7805.

[232] A. Chowdhury and G. Boons, *Tetrahedron Lett.* **2005**, *46*, 1675.

[233] U. Jonsson, L. Fagerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Lofas, B. Persson, H. Roos, I. Ronnberg, S. Sjolander, E. Stenberg, R. Stahlberg, C. Urbaniczky, H. Ostlin and M. Malmqvist, *Biotechniques* **1991**, *11*, 620.

[234] M. Cannon, G. Papalia, I. Navratilova, R. Fisher, L. Roberts, K. Worthy, A. Stephen, G. Marchesini, E. Collins, D. Casper, H. Qiu, D. Satpaev, S. Liparoto, D. Rice, I. Gorshkova, R. Darling, D. Bennett, M. Sekar, E. Hommema, A. Liang, E. Day, J. Inman, S. Karlicek, S. Ullrich, D. Hodges, T. Chu, E. Sullivan, J. Simpson, A. Rafique, B. Luginbuhl, S. Westin, M. Bynum, P. Cachia, Y. Li, D. Kao, A. Neurauter, M. Wong, M. Swanson and D. Myszka, *Anal. Biochem.* **2004**, *330*, 98.

[235] R. J. Guan, A. Roychowdhury, B. Ember, S. Kumar, G. Boons and R. Mariuzza, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17168.

[236] T. Kaneko, W. Goldman, P. Mellroth, H. Steiner, K. Fukase, S. Kusumoto, W. Harley, A. Fox, D. Golenbock and N. Silverman, *Immunity* **2004**, *20*, 637.

[237] S. Girardin, L. Travassos, M. Herve, D. Blanot, I. Boneca, D. Philpott, P. Sansonetti and D. Mengin-Lecreulx, *J. Biol. Chem.* **2003**, *278*, 41702.

[238] S. Girardin, I. Boneca, L. Carneiro, A. Antignac, M. Jehanno, J. Viala, K. Tedin, M. Taha, A. Labigne, U. Zahringer, A. Coyle, J. Bertin, P. Sansonetti and D. Philpott, *Science* **2003**, *300*, 1584.

[239] M. Chamaillard, M. Hashimoto, Y. Horie, J. Masumoto, S. Qiu, L. Saab, Y. Ogura, A. Kawasaki, K. Fukase, S. Kusumoto, M. Valvano, S. Foster, T. Mak, G. Nunez and N. Inohara, *Nat. Immunol.* **2003**, *4*, 702.