Molecular basis for galectin-ligand interactions

Design, synthesis and analysis of ligands

Peterson, Kristoffer

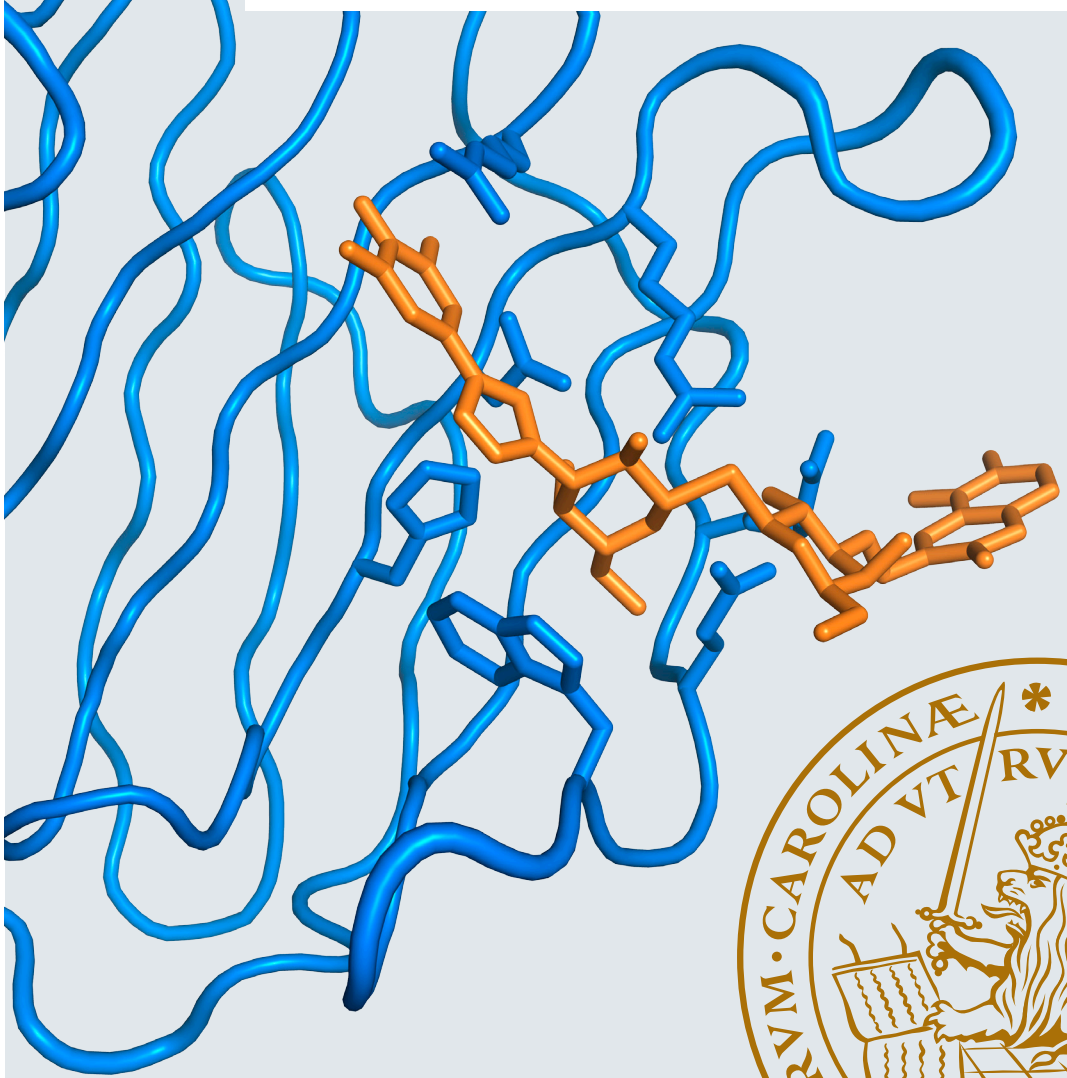
2018

Document Version:
Other version

Molecular basis for galectin-ligand interactions

Design, synthesis and analysis of ligands

KRISTOFFER PETERSON | CENTRE FOR ANALYSIS AND SYNTHESIS | LUND UNIVERSITY



Molecular basis for galectin-ligand interactions

Design, synthesis and analysis of ligands

Kristoffer Peterson

LUND
UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Science, Lund University, Sweden.
To be defended in lecture hall K:C, Kemicentrum,
on Friday 9th of March 2018 at 9:30 a.m.

Faculty opponent
Professor Morten Grötli
Department of Chemistry & Molecular Biology, Gothenburg University, Sweden

Organization	Document name
LUND UNIVERSITY	DOCTORAL DISSERTATION
	Date of issue 2018-02-13
Author: Kristoffer Peterson	Sponsoring organization

Title and subtitle

Molecular basis for galectin-ligand interactions: Design, synthesis and analysis of ligands

Abstract

Galectins are a class of β-galactoside-binding proteins that bind glycoconjugates and have been implicated in cancer, regulation of immunity and inflammation. Design and synthesis have achieved highly potent and selective galectin ligands that can inhibit interactions with glycoproteins and have consequent cellular effects. These ligands can be used as tools to further elucidate the roles of galectins in biological processes, and also, potentially, to diagnose and treat diseases. The present theis is about further development of such galectin ligands. A more robust synthetic route to 3-azido-3-deoxy-β-D-galactopyranosides, key intermediates in the synthesis of previous galectin ligands, has been developed. Bis-3-(4-aryl-1,2,3-triazol-1-yl)-thiodigalactosides are potent galectin-1 and galectin-3 ligands and by screening different aryl groups, the affinities and selectivities for galectin-1 and galectin-3 were improved. In case of galectin-1, the aryl group binds in a smaller binding pocket than in galectin-3, thus five-membered heterocycles were screened with the 2-thiazole having the highest galectin-1 affinity. In case of galectin-3, substituted phenyls were screened with the 3,4,5-trifluorophenyl having the highest galectin-3 affinity. Introducing these aryl groups onto thiodigalactosides resulted in ligands with single-digit nM affinity and 10-50 fold selectivity towards either of galectin-1 or galectin-3. Structural analysis of the galectin-3 ligands identified orthogonal multipolar fluorine-amide interactions and cation-π interactions as main contributors to the high affinity. Based on these findings, monosaccharide derivatives with high selectivity and low nM galectin-3 affinities were developed. Galectin-3 is a biomarker used to diagnose heart failure and through immobilization of a highly potent galectin-3 ligand in a microtiter plate, an assay has been developed that binds both intact and truncated galectin-3 C-terminal domain.

Key words

Galectin, structure-based design, thiodigalactoside, Huisgen 1,3-dipolar cycloaddition, protein-ligand binding interactions

Classification system and/or index terms (if any)

Supplementary bibliographical information		Language English
ISSN and key title		ISBN 978-91-7422-565-5 (print) ISBN 978-91-7422-566-2 (pdf)
Recipient's notes	Number of pages 148	Price
	Security classification	

Molecular basis for galectin-ligand interactions

Design, synthesis and analysis of ligands

Kristoffer Peterson

LUND
UNIVERSITY

Cover by Kristoffer Peterson

Faculty of Science
Department of Chemistry

ISBN 978-91-7422-765-5 (print)
ISBN 978-91-7422-766-2 (pdf)

Printed in Sweden by Media-Tryck, Lund University
Lund 2018

“A person who never made a mistake never tried anything new”

-Albert Einstein

Populärvetenskaplig sammanfattning

I vår kropp finns en myriad av proteiner och celler med viktiga funktioner och de flesta av dem är täckta av kolhydrater. Galektiner är en familj proteiner som har en ficka där olika kolhydrater kan binda in. Genom att kolhydrater på celler eller andra proteiner binder in till denna ficka kan galektiner påverka cellulära funktioner. Vid cancer och inflammation har överproduktion av vissa galektiner observerats, vilket kan vara en indikation på att dessa sjukdomar kan behandlas genom att stoppa bindningen mellan galektiner och proteiner. Detta kan uppnås med hjälp av syntetiska hämmarmolekyler som blockerar den kolhydratbindande fickan hos galektiner.

För att designa en passande hämmarmolekyl behövs information om de molekylära strukturerna i bindningsfickan. Ett sätt att erhålla denna information är att studera kristallstrukturer av naturliga eller syntetiska hämmarmolekyler bundna till galektiner, detta kan sedan användas för att generera idéer till design av mer potenta galektinhämmare. Potenta galektinhämmare kan användas för att studera biologiska processer, biofysikaliska drivkrafter bakom protein-ligand inbindning och fungera som preparat för att diagnostisera och behandla sjukdomar.

I denna avhandling har jag tillverkat hämmarmolekyler för att systematiskt studera och förbättra inbindningen till galektin-1 respektive galektin-3, vilket har resulterat i ett flertal högpotenta och selektiva hämmare. I min jakt på nya högpotenta galektinhämmare utvecklades även en ny syntesmetod för att möjliggöra storskalig syntes av en viktig intermediär molekyl.

Galektin-3 används idag som biomarkör för att diagnostisera hjärtfel men nuvarande metoder med antikroppar kan inte identifiera alla former av galektin-3. Våra utvecklade hämmare binder bra till alla former av galektin-3 så vi har fäst upp en starkt bindande galektinhämmare på en analysplatta som förhoppningsvis kan förbättra nuvarande analysmetoder.

Abbreviations

5-FAM alkyne	5-carboxyfluorescein, propargylamide
5-FAM-NHS	5-carboxyfluorescein, succinimidyl ester
ADME	absorption, distribution, metabolism, excretion
A_{max}	maximum anisotropy
CRD	carbohydrate recognition domain
DCM	dichloromethane
DIPEA	diisopropylethylamine
FP	fluorescence polarization
Imz	imidazylate
ITC	isothermal titration calorimetry
K_d	dissociation constant
LacNAc	*N*-acetyllactosamine
LLE	ligand lipophilicity efficiency
LogP	logarithm of the partition-coefficient
PSA	polar surface area
rt	room temperature
SEM	standard error of the mean
TBAF	tetrabutylammonium fluoride
TDG	unsubstituted thiodigalactoside
Tf	triflate
TIPS	triisopropylsilane

Abbreviations found in the ACS Style Guide are not included.

List of papers

This thesis summarizes and supplements the following papers, referred to by their Roman numerals. Paper I is reprinted with kind permission of Taylor & Francis. Paper II is reprinted with kind permission of American Chemical Society. Paper VI is reprinted with kind permission of Wiley.

I. Peterson, K.; Weymouth-Wilson, A.; Nilsson, U. J. Aryl sulfonates in inversions at secondary carbohydrate hydroxyls: a new and improved route towards 3-azido-3-deoxy-β-D-galactopyranosides. *J. Carbohydr. Chem.* **2015**, *34*, 490-499.

Contribution: Performed all experimental work and contributed to writing the manuscript.

II. Peterson, K.; Kumar, R.; Stenström, O.; Verma, P.; Verma, P. R.; Håkansson, M.; Kahl-Knutsson, B.; Zetterberg, F.; Leffler, H.; Akke, M.; Logan, D. T.; Nilsson, U. J. Systematic tuning of fluoro-galectin-3 interactions provides thiodigalactoside derivatives with single digit nM affinity and high selectivity. *J. Med. Chem.* pre-print ahead of press, **2018**, doi: 10.1021/acs.jmedchem.7b01626

Contribution: Performed the synthesis of all compounds, helped design the ITC experiments and contributed to writing the manuscript.

III. Peterson, K.; Collins, P.; Kahl-Knutsson, B.; Zetterberg, F.; Leffler, H.; Blanchard, H.; Nilsson, U. J. Aromatic heterocycle galectin-1 interactions for single-digit nM affinity ligands. *Manuscript*.

Contribution: Performed the synthesis of all compounds and contributed to writing the manuscript.

IV. Peterson, K.; Kumar, R.; Misini Ignjatović, M.; Leffler, H.; Ryde, U.; Logan, D. T.; Nilsson, U. J. Polyfluoroaryl substituents influencing galectin-3 binding through steric- and electronic conjugation effects. *Manuscript.*

Contribution: Performed the synthesis of all compounds and contributed to writing the manuscript.

V. Kumar, R.; Peterson, K.; Olsson, M.; Wallerstein, J.; Leffler, H.; Akke, M.; Ryde, U.; Nilsson, U. J.; Logan, D. T. Structure and thermodynamics of ligand-fluorine interactions with galectin-3 backbone- and side-chain amides. *Manuscript*.

Contribution: Performed the synthesis of all compounds and contributed to writing the manuscript.

VI. Zetterberg, F. R.; Peterson, K.; Johnsson, R. E.; Brimert, T.; Håkansson, M.; Logan, D. T.; Leffler, H.; Nilsson, U. J. Monosaccharide derivatives with low nM lectin affinity and high selectivity based on combined fluorine-amide, phenyl-arginine, sulfur-π, and halogen bond interactions. *ChemMedChem* **2018**, *13*, 133-137.

Contribution: Contributed to ligand analysis and writing the manuscript.

VII. Peterson, K.; Stegmayr, J.; Karlsson, A.; Leffler, H.; Nilsson, U. J. A galectin-3 and galectin-3C assay based on a high-affinity small-molecule ligand immobilized in microwells. *Manuscript*.

Contribution: Performed the synthesis of all compounds and contributed to writing the manuscript.

List of papers not included in this thesis

VIII. Uchino, Y.; Woodward, A. M.; Mauris, J.; Peterson, K.; Verma, P.; Nilsson, U. J.; Leffler, H.; Rajaiya, J.; Argüeso, P. Galectin-3 is an amplifier of the interleukin-1β-mediated inflammatory response in corneal keratinocytes. *Immunology*, **2018**. *Accepted*.

Content

1. Hypotheses, aims and outline

Over the years the importance of galectins (β-galactoside-binding proteins) in biological processes have become more apparent. In order to study and further elucidate the roles of galectins in biological systems, different kinds of galectin ligands have been developed, and some have emerged as drug candidates.

I proposed that synthetic novel small-molecule galectin ligands would aid in improving our understanding of the galectins' role in biological events and provide valuable tools in diagnosing and treating diseases. Interactions with ligand-fluorines in a subsite of the galectin-3 binding pocket have been reported to enhance the galectin-3 binding of thiodigalactosides substituted with fluorinated phenyltriazoles. I hypothesized that the potency and selectivity of galectin-3 ligands can be improved through optimization of these fluorine interactions. The corresponding subsite in galectin-1 is smaller and the use of a thiophene instead of fluorinated phenyls has proven to favor galectin-1 binding. I hypothesized that by screening various five-membered heterocycles the potency and selectivity of galectin-1 ligands can be improved. Most natural and synthetic galectin ligands are based on disaccharides. During my investigation of ligand-fluorine interactions with galectin-3 I hypothesized that replacing one of the galactose units in thiodigalactosides with an optimized thio-arene would result in potent monosaccharide-based galectin-3 ligands.

The first aim of this thesis was to design and synthesize galactosides substituted with various aryltriazoles, and evaluate their galectin-1 and galectin-3 affinities. The second aim was to study galectin-3-ligand binding interactions and gain a better understanding of how galectin-1 and galectin-3 ligands bind.

The focus of this thesis is to improve and decipher the mechanism behind binding of synthesized galectin-1 and galectin-3 ligands. Instead of describing the different projects individually I have tried to describe everything as one big project. In chapter 2 the reader is given a short introduction to carbohydrates, galectins and protein-ligand interactions. Chapter 3 starts off with a general introduction to medicinal chemistry followed by the synthesis of all ligands. The evaluation of the synthesized ligands is then discussed in terms of affinity and selectivity (Chapter 4) and binding interactions (Chapter 5). Chapter 6 discusses galectin ligands as a diagnostic tool for inflammatory conditions.

2. Carbohydrates, galectins and how they fit together

Carbohydrates are the most abundant organic molecules found on earth, with roles ranging from structural material in plants (cellulose) to components in our DNA (deoxyribose). Living organisms use their DNA to synthesize proteins that can be further functionalized by the addition of carbohydrates, resulting in glycoproteins. If carbohydrates instead are added to lipids or peptides the results are glycolipids or glycopeptides, and they are together with glycoproteins referred to as glycoconjugates. Glycoconjugates are important in many biological processes, e.g. the immune system, cell growth, cell-cell adhesion and inflammation.[1] With the plethora of biological activity going on in the human body cell-cell communication is crucial and one class of proteins responsible for mediating biological recognition events are lectins, the protein class the galectins belong to.

2.1 Carbohydrates

Carbohydrates are, as the name implies, hydrates of carbon $(CH_2O)_n$. The number of carbohydrate units present determines if a carbohydrate is a mono- (1 unit), di- (2 units), oligo- (2-10 units) or polysaccharide (>10 units). Saccharides can be either acyclic or cyclic molecules but typically exist in their cyclic conformation. The most common monosaccharides have six carbons, termed hexoses or pyranoses (if cyclic), with glucose and galactose being well known examples. Pyranoses are often drawn in their chair conformation to visualize the stereochemistry of the five stereogenic centers. A saccharide is assigned either as an α or β anomer depending on the stereochemistry of the substituent at C1 (anomeric center) in relation to the stereogenic center furthest from the anomeric center. In galactose, the main saccharide discussed in this thesis, α refers to an axial substituent (Figure 1B) at C1 and β an equatorial substituent (Figure 1C). Carbohydrates are chiral compounds that exist as enantiomers and the different enantiomers are denoted D or L depending on if the hydroxyl group furthest from the carbonyl group is on the right (D) or left (L) side when drawn in the Fischer

projection (Figure 1A). Galactose and glucose D enantiomers are only found in nature. Thiodigalactoside is a disaccharide made up of two galactose units linked together by a sulfur in (1→1) fashion. Thiodigalactosides with β stereochemistry (Figure 1D) is, as will be evident in this thesis, a suitable core structure for galectin ligands and the unsubstituted thiodigalactoside (TDG) is often used as a reference compound in biological assays.

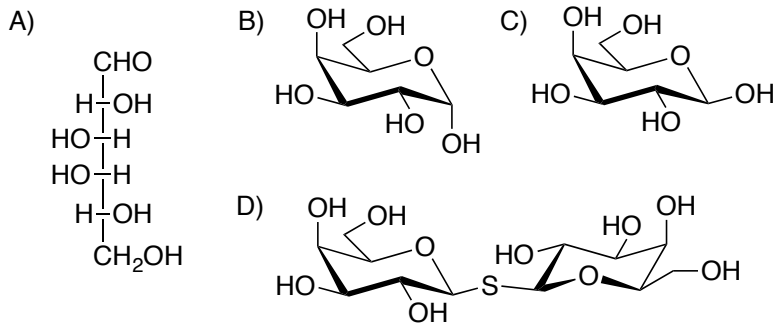


Figure 1. A) Fischer projection of D-galactose. Structures of B) α-D-galactose, C) β-D-galactose and D) TDG.

2.2 Galectins

2.2.1 Background

In 1994 Barondes *et al.*[2] coined the term galectin, and to be included in the galectin family two criteria must be fulfilled: "affinity for β-galactosides and significant sequence similarity in the carbohydrate-binding site". Today there are 15 members of the galectin family, that exists as monomers, dimers or oligomers (depending on conditions such as concentration and presence of ligand). They are divided into three subgroups, proto-, chimera- and tandem-repeat type (Figure 2).[3]

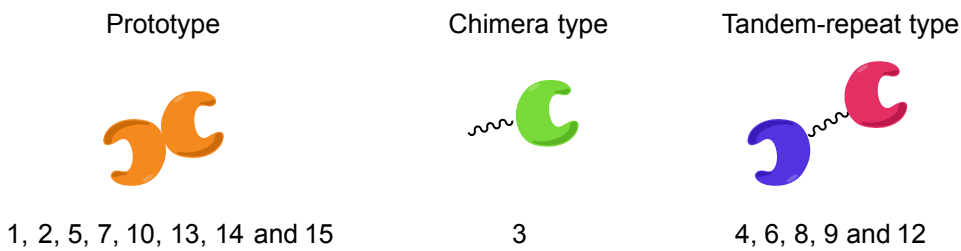


Figure 2. Galectin subgroups showing proto-, chimera- and tandem-repeat type and the galectins belonging to that subgroup (below).

All galectins have a carbohydrate recognition domain (CRD), where carbohydrate-based ligands bind, consisting of about 135 amino acids that is folded into a β-sandwich structure. The two β-sheets form a long groove that is divided into five subsites (Figure 3) where every subsite except E corresponds to a β-strand, with the galactose unit binding in the middle subsite C. The galactose C3 substituent is positioned towards subsites A-B and the C1 substituent is positioned towards subsites D-E.[3]

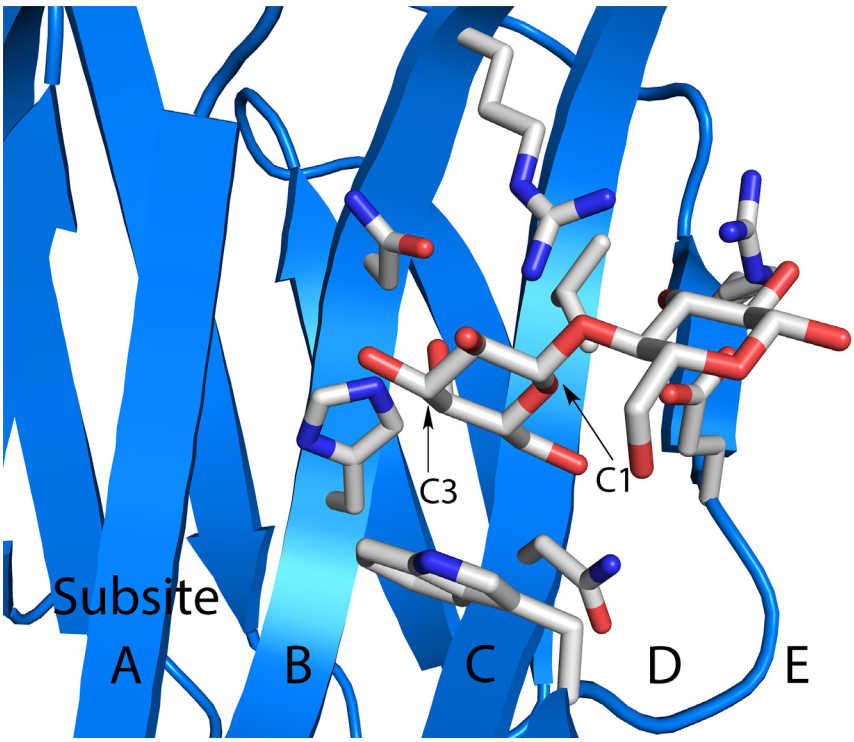


Figure 3. Illustration of the five subsites (A-E), exemplified with lactose bound to the CRD of galectin-3. Carbons C1 and C3 of the galactose residue are indicated. The amino acids shown as sticks are present in both galectin-1 and galectin-3.

The CRD of galectin-3 is located at the C-terminal while the N-terminal has a non-carbohydrate binding domain, which is believed to be responsible for the oligomerization[4] of galectin-3 at higher concentrations and thus allowing multivalent interactions with glycoconjugates. Oligomerization can be a problem in *in vitro* studies such as isothermal titration calorimetry and X-ray crystallography. To avoid oligomerization a truncated version of galectin-3 only consisting of the CRD (galectin-3C) is often used.[5]

2.2.2 Biological location and function

Galectins are present throughout the body and function both extra- and intracellularly. Galectin-1 is found in most tissue, while galectin-3 is mainly found in macrophages, epithelia and tumors.[6] The cell surfaces of human cells are coated with glycoconjugates that in some cases are galactose-containing oligosaccharides suitable for interactions with the CRD of galectins.[7] As most galectins are either bi- or multivalent they can cross-link glycoconjugates and form lattices, and thereby modulating these glycoconjugates localization, transport and residence times at cell surfaces or in intracellular vesicles. This lattice formation will have different effects at the cellular level depending on the galectin type and the glycoconjugates present.[8] There are numerous biological events where galectins have been implicated to play important roles, mainly concerning cancer,[9,10] the regulation of immunity[11] and inflammation[11,12].

2.2.3 Ligands

As mentioned earlier, part of the definition of a galectin is the affinity for β-galactosides. β-Galactosides are of course found in galactose derivatives, but the most common reference compounds found in literature are the disaccharides lactose, *N*-acetyllactosamine (LacNAc) and TDG. On cell surfaces LacNAc-branches are typically found on glycoconjugates that interact with galectins. Galectin ligands can be divided into mono- and multivalent ligands depending on how many galectin CRD they can bind. Although multivalency can improve inhibitory potency[13], only monovalent galectin ligands will be considered in this thesis.

There are several benefits of using small synthetic molecules as galectin ligands, compared to natural saccharides. Important aspects in development of lead compounds are affinity and selectivity, and through synthetic modification of saccharides like lactose and TDG, selective ligands with high affinity towards a specific galectin can be obtained. Synthetic ligands also tend to have lower polarity and less hydrogen bonding groups compared to natural saccharides. Previously reported ligands towards galectin-1 and galectin-3 (a selection is presented in Figure 4) are usually based on 1- and 3-substitutions of galactose[14–16], 3-substitution of lactose[17–21] or 3,3'-disubstitutions of thiodigalactoside[22–27]. Structural analyses (Figure 5) of galectin-1 and galectin-3 in complex with TDG also show that these positions are suitable for modifications. However, modifications at C2 of galactose[28] and talose[29] have been reported, while 4-OH and 6-OH are found to be important to galectin binding.

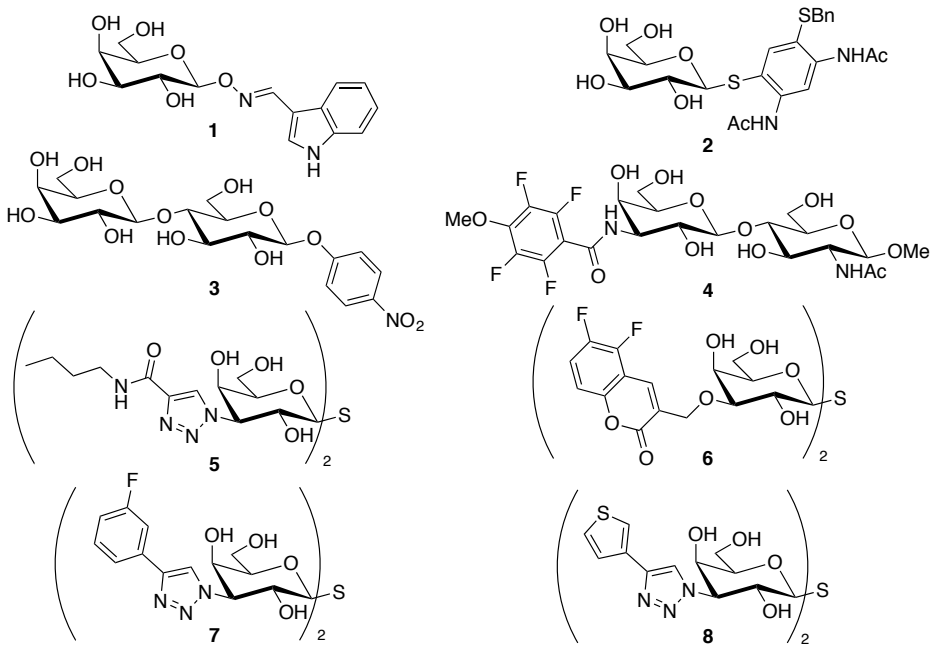


Figure 4. A selection of previously reported galectin ligands based on galactose (**1**[16] and **2**[30]), lactose (**3**[20]), LacNAc (**4**[31]) and thiodigalactoside (**5**[25], **6**[32], **7**[27] and **8**[27]).

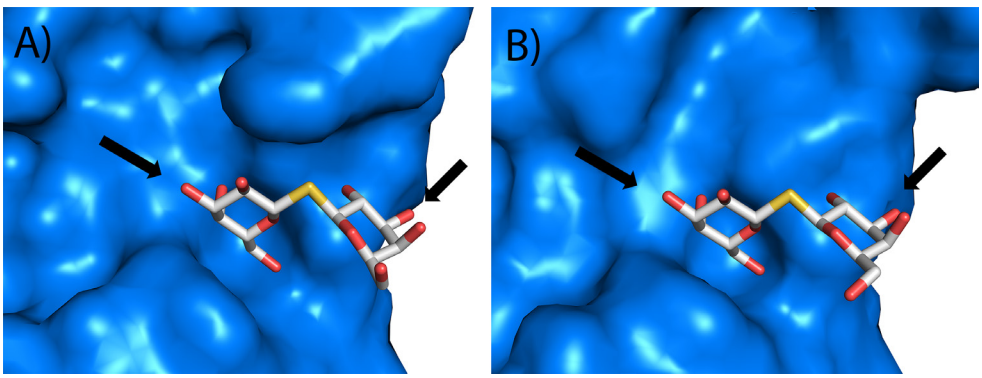


Figure 5. Galectin-1 (A) and galectin-3C (B) in complex with TDG with possible 3,3'-substitutions indicated.

2.3 Protein-ligand binding interactions

In chapter 3, ligand design will be discussed, but first we need to know what drives a ligand to bind to a protein. Simply put, a ligand binds to a protein if the Gibbs free energy (ΔG), *i.e.* the sum of enthalpic (ΔH) and entropic (-TΔS) contributions, is negative.[33] The thermodynamic fingerprint of a protein-ligand binding event is comprised of intermolecular interactions, solvation processes and conformational changes. The intermolecular interactions of a protein-ligand complex can be deduced from available crystal structures, while analysis of solvation processes and conformational changes are more difficult to deduce, because they are dynamic processes and the crystal structure is a static image. The crystal structure shows the electron density of the protein-ligand complex, meaning hydrogens with only one electron are usually not detected.[33]

In structure-based ligand design (see section 3.1.1) intermolecular interactions such as dipole-dipole interactions, hydrophobic interactions and interactions involving π-systems are optimized. A dipole arises in a bond when one atom is more electronegative than the other resulting in one partially negative and one partially positive atom. When the partially negative atom of one dipole is positioned close to a partially positive atom of another dipole a dipole-dipole interaction is formed. A well-known dipole-dipole interaction is the hydrogen bond where the hydrogen of typically a NH or OH group interacts with an electronegative atom like oxygen or nitrogen. In hydrogen bonds, the angle and distance between participating atoms are important for a favorable orbital overlap and the donor−hydrogen···acceptor angle generally is above 150°, while the distance varies between 2.7-3.0 Å.[33]

Organofluorine compounds are common in drug development and aside from increased hydrophobicity the highly polarized C-F bond can form dipole-dipole interactions with moieties in the protein that have a partial positive charge, like amides. Interactions between fluorines and backbone amides are known[34,35] and have been systematically studied[36], while fluorine interactions with side-chain amides are harder to find in literature. In the multipolar C−F···C=O interaction with backbone amides there is clear preference for the fluorine to be positioned orthogonal to the carbon at distances ≤ 3.0 Å, while at longer distances the angle (denoted *a* in Figure 6A) dependence is weaker. Because of this geometric preference, the interaction is sometimes referred to as an orthogonal multipolar fluorine-amide interaction.[35] Measurements of the orthogonal multipolar fluorine-amide interaction in nonpolar solvents have found an attractive interaction of 0.8-1.5 kJ mol^{-1}.[36] In a PDB search in 2007[34] several fluorine interactions with side-chain amides were found. Based on a statistical analysis the angle (denoted a_1 and a_2 in Figure 6B) dependence is weaker in fluorine interactions with side-chain

amides than with backbone amides. The shortest distance in the fluorine interactions with side-chain amides was found to be F···N-C at ≤ 3.1 Å.

Heavier halogens than fluorine have a σ-hole[37], that is a positive electrostatic potential opposite the C-X bond and its size increases further down in the periodic table. Carbonyl oxygens, or other hydrogen bond acceptors, can form attractive interactions with the σ-hole of halogens in a preferred linear C−X···acceptor arrangement resulting in a halogen bond (Figure 6C).[37]

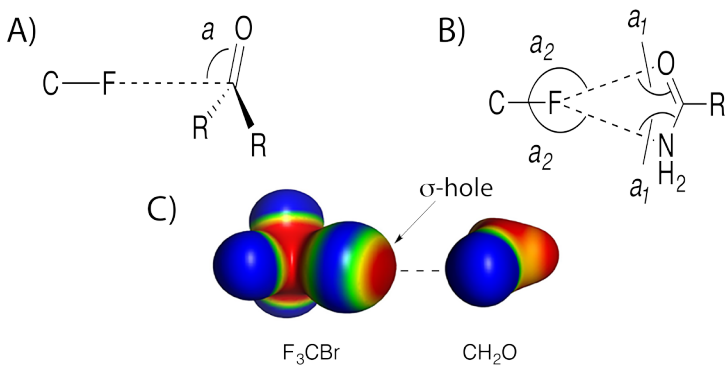


Figure 6. A) Overview of a fluorine-amide interaction with A) backbone amide and B) side-chain amide. C) Overview of a halogen bond between a carbonyl oxygen and the σ-hole of a C−X bond. Illustrated with F_3CBr and CH_2O where high electron density is shown in blue and low electrondensity in red.

Interactions involving π-systems include for example π-stacking, cation-π and sulfur-π interactions[38]. They are important in biomolecular recognition and protein folding due to their ubiquitous binding partners in proteins, which have four aryl-containing amino acids (Trp, Phe, Tyr and His), two cation-containing amino acids (Arg and Lys) and two sulfur-containing amino acids (Met and Cys). In π-stacking (Figure 7A) the two π-systems are preferentially oriented edge-to-face (T-shaped), eclipsed face-to-face or parallel displaced to maximize the interaction. The electrons in a π-system can form cation-π interactions (Figure 7B) through electrostatic attraction with cations and the strength of the interaction depends on the nature of the π-system (*e.g.* presence of heteroatoms) and its substituents. Electron-withdrawing groups, like fluorines, will weaken the cation-π interaction, while electron-releasing groups will strengthen this interaction. Sulfur-π interactions (Figure 7C) are formed when a lone pair on sulfur approaches either atop the face of the aryl ring in a S···π arrangement or at the edge of the aryl ring in a S···HC(π) arrangement.[38] Sulfur has regions of positive electrostatic potential that can be compared with the σ-hole of halogens. The σ-holes of sulfur have an

orthogonal orientation with regard to the lone pairs, and may interact favorably with π-systems in a S···π arrangement.[39]

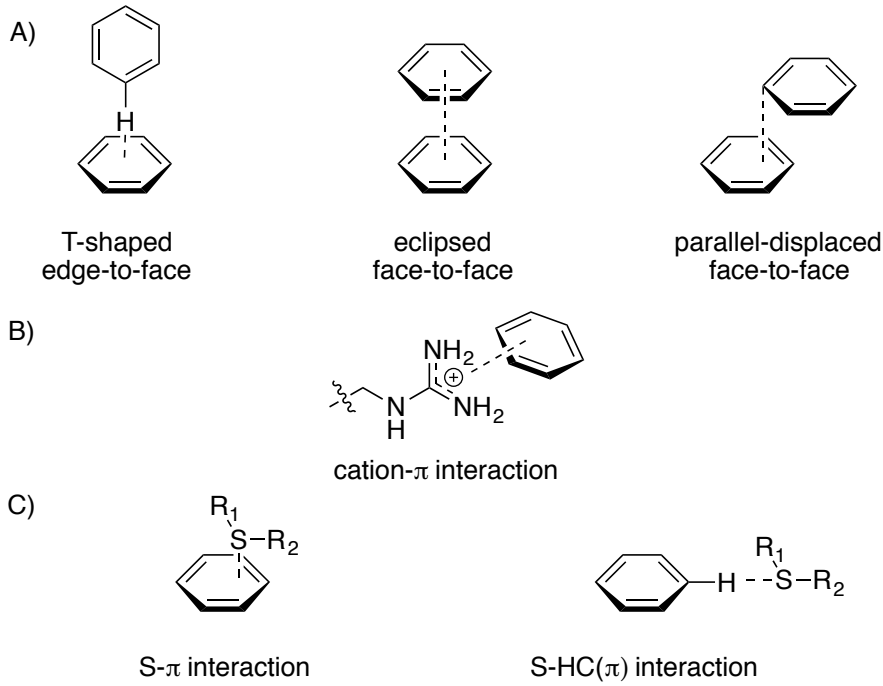


Figure 7. Overview of the different interactions involving π-systems. A) π-stacking, B) cation-π interaction and C) sulfur-π interactions.

3. Design and synthesis of galectin ligands

3.1 Medicinal chemistry

The first efforts to isolate, purify and determine the structure of the active substances in naturally occurring drugs took place in the middle of the nineteenth century and it paved the way for chemists to make synthetic versions and analogs. In the early years of drug discovery a trial and error approach was used to discover new drugs, but over the years patterns and strategies emerged. Today these strategies have evolved, much as a result of advances in structural biology and our understanding of biological function at the cellular and molecular level. Lead compounds are still, in many cases, based on natural ligands but structural knowledge of the biological target has greatly helped the design of new drugs, and technological advancements, such as computer modeling and high throughput screening have helped the generation of promising drug candidates.[40]

3.1.1 Ligand design

The analysis of lead compounds to determine which parts of a molecule are important for biological activity can be used to establish the structure-activity relationship (SAR). New ligands can be designed (ligand-based design), synthesized and evaluated to give an updated SAR. This can be repeated until ligands with satisfactory properties are obtained. Another way to design new ligands is to study the three-dimensional structure of the biological target and identify possible binding interactions (structure-based design).[40] In fragment-based design, small molecules (fragments) are evaluated towards the target to identify fragments with high affinity relative to its size. The information from the fragments is then combined to design bigger, hopefully better, ligands. The strategies used in drug development can also be utilized to design molecular tools for studying biological events and to elucidate protein-ligand binding interactions. Since the design of new ligands is an iterative process it is essential to have reliable methods for ligand evaluation. There are several techniques[41] available for

evaluating protein-ligand binding and for the work presented in this thesis fluorescence polarization (FP) and isothermal titration calorimetry (ITC) have been used to determine dissociation constants (K_d), while FP, ITC and X-ray crystallography have been used to study galectin-3 binding interactions. These techniques will be covered in more detail in chapters 4 and 5.

3.1.2 Scoring systems

In past decades the importance of drug metabolism and pharmacokinetics has become apparent and determination of physiochemical properties along with *in vitro* and *in vivo* profiling are commonplace in drug development.[42] An in-depth analysis of ADME, pharmacokinetics and drug metabolism are not within the scope of this thesis, instead a brief overview of the scoring systems and parameters I find relevant are given.

A good drug candidate does not only have to exhibit high potency it also has to be able to reach its intended target. When formulating a drug, different properties are considered depending on how the drug is going to be administered *e.g.* orally, intravenously or as an inhaled aerosol. In small-molecule drug discovery the aim is generally to find orally available drugs, which means it needs to have good intestinal uptake (*i.e.* high membrane permeability). Since it is expensive and labor-intensive to test membrane permeability for every synthesized compound, Lipinski *et al.*[43] in 1997 identified several easily measured properties to predict membrane permeability, which was coined the "rule of 5". The "rule of 5" is based on data from of about 50 000 drugs from the Derwent Co and it does not include compound classes that are substrates for biological transporters. The "rule of 5" can be viewed as a guideline, and states that compounds with >5 H-bond donors, >500 in molecular weight, >5 in LogP and >10 H-bond acceptors are unlikely to permeate membranes.

Since the introduction of the "rule of 5" other scoring systems to evaluate drug candidates have emerged, *e.g.* polar surface area (PSA)[44] and ligand lipophilic efficiency (LLE)[45]. PSA is used in predicting membrane permeability and it is defined as the surface area occupied by nitrogens and oxygens and their polar hydrogens. The basis is that less polar compounds have better membrane permeability and based on measurements of drugs with the Caco-2 system[46], drugs with a PSA value greater than 140 $Å^2$ will have poor membrane permeability[44].

Lipophilicity is an important property of drug candidates and because of desolvation, an increase in lipophilicity often leads to an increase in binding affinity. However, high lipophilicity is correlated with promiscuity and lack of target specificity, potentially leading to toxic side effects. Because our bodies are equipped to get rid of lipophilic compounds, *e.g.* cytochrome P450, high

lipophilicity usually means high metabolic instability. The concept of LLE is to estimate (LLE = pK_d – LogP) binding efficiency with respect to ligand physical properties such as size and lipophilicity.[45]

3.1.3 Synthesis strategies

The discussion so far has been about how to design ligands and what properties they should have but not how to make them, which is an important aspect in drug development. Most people who have performed organic synthesis would agree that more time is spent on synthesizing compounds than on designing them, which is why it is important to prioritize design ideas. Strategies for reducing the synthesis workload include divergent synthesis that uses key intermediates from which series of ligands are synthesized. A good way to generate ligand series from a common intermediate is to utilize robust and versatile reactions that incorporate structural motifs and allow substituent variations. This can be achieved with "click chemistry" a concept introduced in 2001 by Sharpless *et al.*[47]. Click reactions are high yielding, wide in scope, stereospecific and ideally insensitive to oxygen and water, which typically are achieved by having a high thermodynamic driving force of > 80 kJ mol^{-1}. Click reactions tend to be highly selective for a single product and many carbon-heteroatom bond-forming reaction classes are examples of click reactions, especially the Cu(I)-catalyzed Huisgen 1,3-dipolar cycloaddition[48,49] (Figure 8) between an azide and an alkyne.

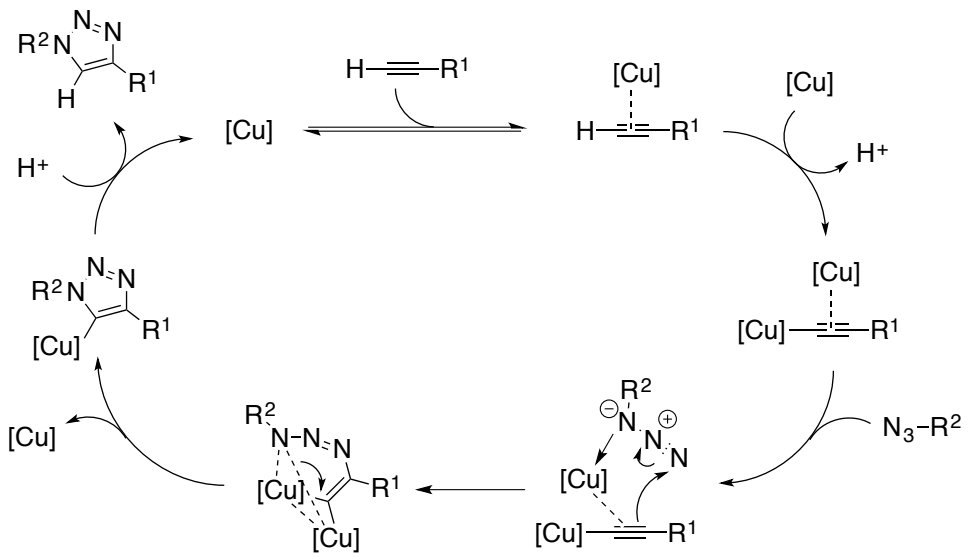
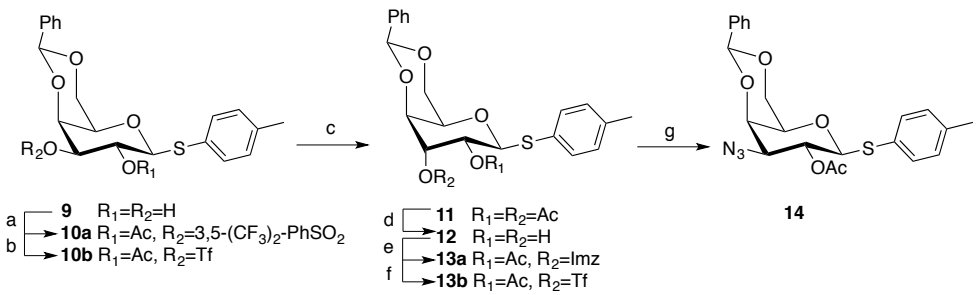


Figure 8. Proposed[50] catalytic cycle of the Cu(I)-catalyzed Huisgen 1,3-dipolar cycloaddition.

3.2 Synthesis of key intermediates (Paper I)

Over the years, several galectin ligands with either a triazole or an amide at galactose C3 have been reported (see section 2.2.3), and both can be made from a common functional group, an azide. Triazoles can be made with click chemistry from azides, while reduction of the azide to an amine enables amide couplings. The synthesis of the versatile 3-azido-3-deoxy-β-D-galactopyranosides key intermediates has been reported by Öberg *et al.*[51] using a double inversion route with triflates as leaving group. One drawback of this route was the instability of the triflate intermediates that, due to early onset of exothermic decomposition, compromised scale-up. Since we wanted to be able to make 3-azido-3-deoxy-β-D-galactopyranosides on large scale, we decided to investigate if the triflate leaving group could be replaced with more stable aryl sulfonates. The challenge is that stability usually means less reactivity. In order to achieve good reactivity the triflate was replaced with a 3,5-bis(trifluoromethyl)benzenesulfonate[52] in the first inversion and with an imidazylate[53] in the second (Scheme 1). This led to 3-azido-3-deoxy-β-D-galactopyranoside **14** in yields comparable to those obtained by Öberg *et al.* Measurement of the onset of exothermic decomposition using differential scanning calorimetry did indeed show improved stability of aryl sulfonates **10a** (143 °C) and **13a** (148 °C) compared to the corresponding triflates **10b** (112 °C) and **13b** (110 °C). The increased stability of the aryl sulfonates, compared to the triflates, makes them more suitable in large-scale synthesis of 3-azido-3-deoxy-β-D-galactopyranosides.



Scheme 1. Synthesis of 3-azido-3-deoxy-β-D-galactopyranosides **14**. Reagents and conditions: a) i. 3,5-bis(trifluoromethyl)benzenesulfonyl chloride, 1,2,2,6,6-pentamethylpiperidine, Bu_2SnCl_2, THF, rt; ii. Ac_2O, py, rt; b) i. Tf_2O, py, DCM, rt; ii. AcCl, py, rt, 56 %; c) CsOAc, DMSO, 90 °C, 72 % from **9** via **10a** and 7 % **10a** recovered; d) NaOMe, MeOH, rt, 99 %; e) i. AcCl, py, DCM, 0 °C; ii. SO_2Cl_2, imidazole, DMF, rt; f) i. AcCl, py, DCM, 0 °C; ii. Tf_2O, py, DCM, 0 °C, 66 %; g) Bu_4NN_3, DMSO, 70 °C, 29 % from **12** via **13a**.

3.3 Synthesis of galectin-1 and galectin-3 ligands (Papers II-IV)

3.3.1 Structure of subsites A-B in galectin-1 and galectin-3

The design of the galectin-1 and galectin-3 ligands presented later in this chapter was based on the crystal structures of thiodigalactoside **8** (see Figure 4 for molecular structures of **7-8**) in complex with galectin-1[Paper III] (Figure 9) and thiodigalactoside **7** in complex with galectin-3[27] (Figure 10). The main difference between galectin-1 and galectin-3 is found in subsite A where Arg144 in galectin-3 is positioned to form a cation-π interaction with the aryl group of thiodigalactosides **7-8**. A corresponding arginine is absent in galectin-1, and most other galectins. Instead the aryl group makes a close van der Waals interaction with Val31 in galectin-1. The subsite A binding pocket of galectin-1 is also smaller than in galectin-3, which might explain the higher affinity of the five-membered thiophene ring in **8** compared to the six-membered phenyl ring in **7**.

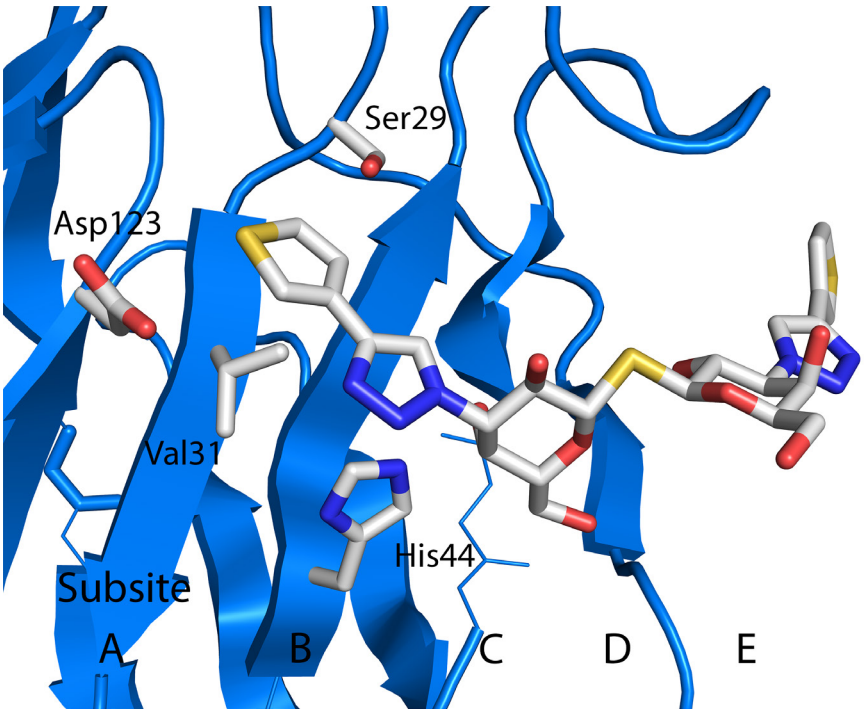


Figure 9. Crystal structure of galectin-1 in complex with thiodigalactoside **8**. The important side chains in subsites A-B are showed as sticks.

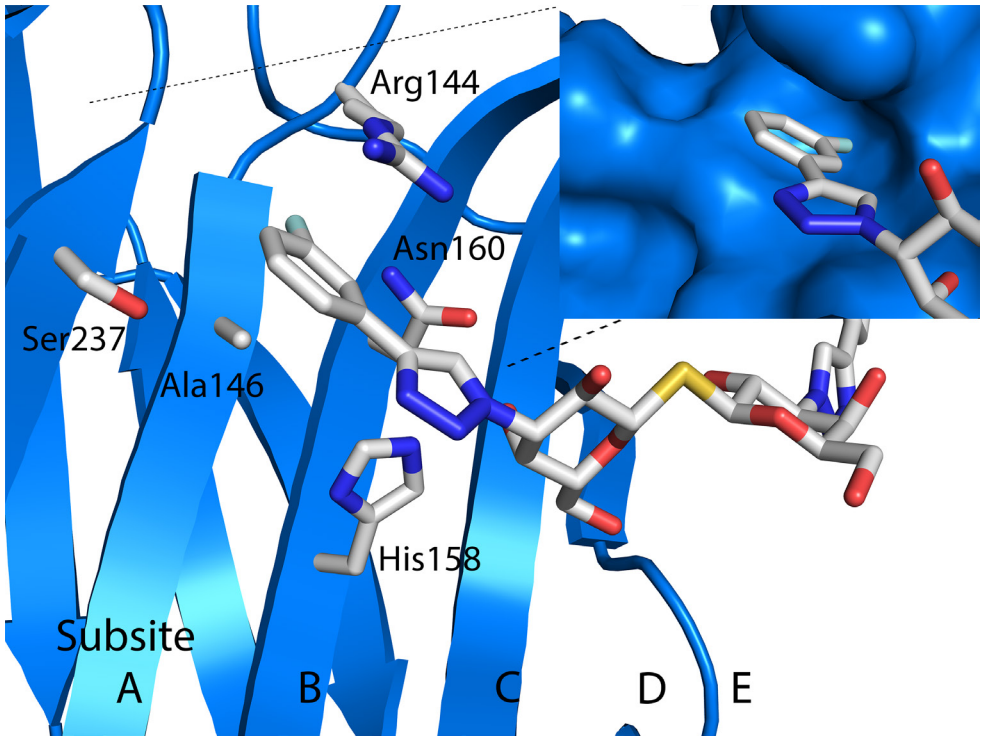
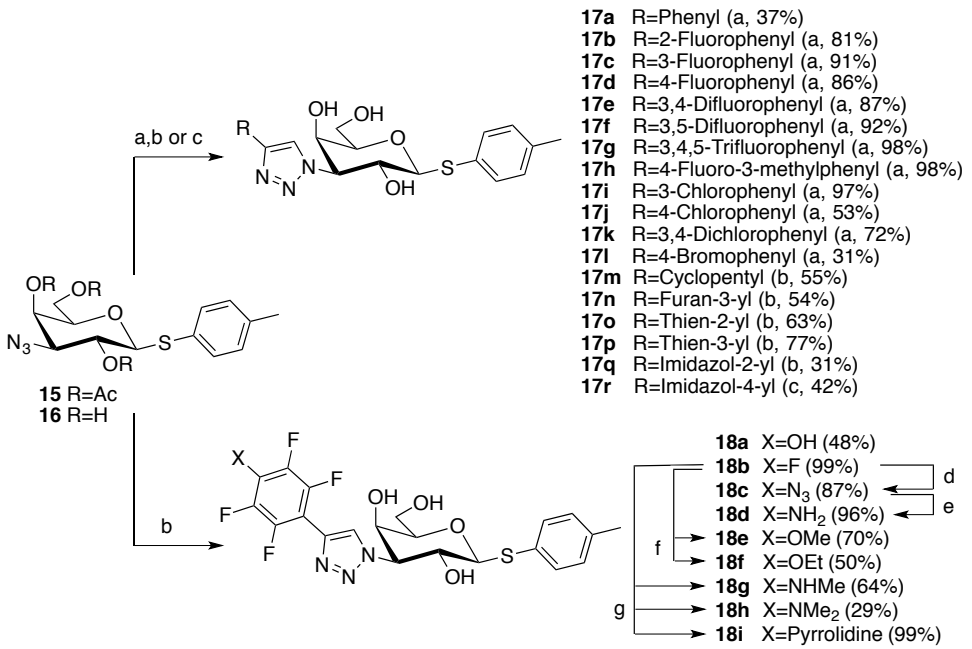


Figure 10. Crystal structures of galectin-3C in complex with thiodigalactoside **7**. The deep pocket in subsite A highlighted with dashed lines. The important side chains in subsites A-B are showed as sticks.

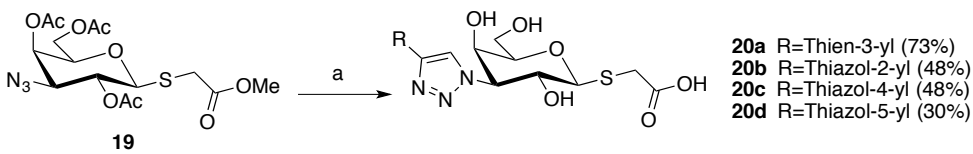
3.3.2 Optimization of binding interactions in subsite A

Two series of 3-(4-aryl-1,2,3-triazol-1-yl)-thiogalactosides (**17a-r** and **18a-i**) were synthesized (Scheme 2) from azides **15** and **16** (obtained by the development of the new double inversion route) using the Cu(I)-catalyzed Huisgen 1,3-dipolar cycloaddition. Thiogalactosides **17a-l**, with a phenyl at the triazole C4, and thiogalactosides **17m-r**, with a five-membered heterocycle at the triazole C4, were designed to optimize affinity towards galectin-3 and galectin-1, respectively. The 2-thiazole analog of triazoles **17** was synthesized but due to poor solubility, even in DMSO-d_6, confirmation of the product could only be achieved with HRMS. Three thiazole analogs were instead synthesized from azide **19**[54] (Scheme 3) to give thiogalactosides **20b-d** equipped with a hydrophilic carboxylic acid substituent at the anomeric position. The 3-thiophene analog (**20a**) to thiazoles **20b-d** was made in order to have a complete series to compare with triazoles **17m-r**.

Penta- and tetrafluorophenyltriazoles **18a-i** were synthesized to study the binding interactions in a deep pocket near Arg144 in galectin-3 (Figure 10) that could potentially accommodate a bigger phenyl substituent. Aromatic nucleophilic substitution of the para fluorine in **18b** with alcohols, amines and NaN_3 gave **18c** and **18e-i**, while substitution with ammonia or hydroxide was not successful. Aromatic nucleophilic substitution was, however, successful by hydroxide on 2,3,4,5,6-pentafluorophenylethynyltrimethylsilane. By using the Cu(I)-catalyzed Huisgen 1,3-dipolar cycloaddition, **18a** was obtained, while azide reduction of **18c** gave amine **18d**.

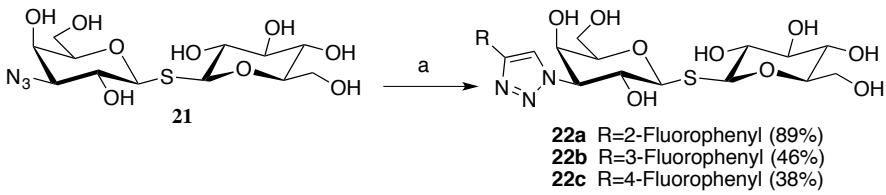


Scheme 2. Synthesis of triazole series **17a-r** and **18a-i**. Reagents and conditions: a) From **15**: i. alkyne, CuI, DIPEA, MeCN, 50 °C; ii. NaOMe, MeOH, rt; b) From **16**: alkyne, CuI, DIPEA, MeCN, 50 °C; c) From **16**: i. 1-{[2-(trimethylsilyl)ethoxy]methyl}-4-[2-(trimethylsilyl)ethynyl]-1*H*-imidazole, CuI, DIPEA, MeCN, 50 °C; ii. BF_3OEt_2, DCM, rt; d) NaN_3, DMF, 60 °C; e) 1,3-propanedithiol, Et_3N, MeOH, rt; f) sodium alkoxide, alcohol, rt; g) amine, K_2CO_3, DMF, 50 °C.



Scheme 3. Synthesis of triazoles **20a-d**. Reagents and conditions: a) i. alkyne, CuI, DIPEA, MeCN, 50 °C; ii. NaOMe, MeOH, rt; iii. LiOH, THF:H_2O (9:1), rt.

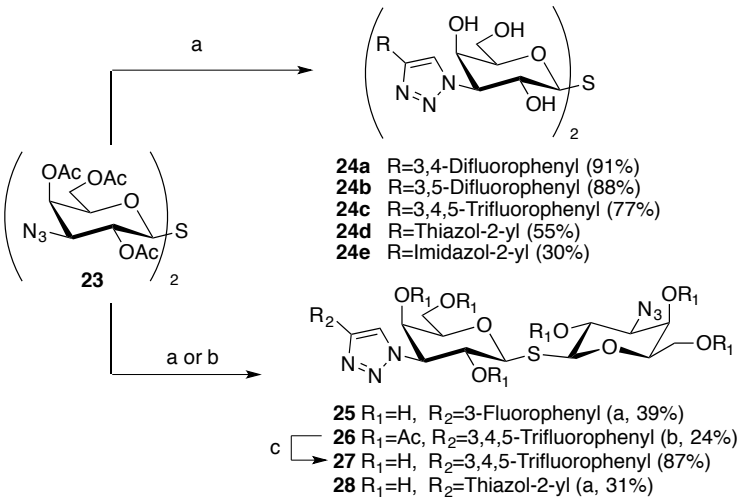
To thoroughly study the contribution of the fluoride-amide interaction (Chapter 5) on phenyltriazoles **17b-d**, the three monofluorinated phenyltriazoles were subjected to ITC analysis. When evaluating ligands with ITC, higher ligand concentrations (H_2O:DMSO, 95:5) than those used in FP are needed, which led to poor solubility of **17b-d**. We hypothesized that replacing the tolyl group in **17b-d** with a polar sugar moiety would increase the water solubility. Glucose was chosen instead of galactose to ensure that the galactose with the triazole moiety bound in subsite C of galectin-3. Hence, galactosylthioglucosides **22a-c** (Scheme 4) were synthesized through click chemistry from azide **21**[55]. Higher water solubility was indeed observed for galactosylthioglucosides **22a-c** as they could be dissolved in water without DMSO for ITC analysis.



Scheme 4. Synthesis of triazoles **22a-c**. Reagents and conditions: a) alkyne, CuI, DIPEA, MeCN, 50 °C.

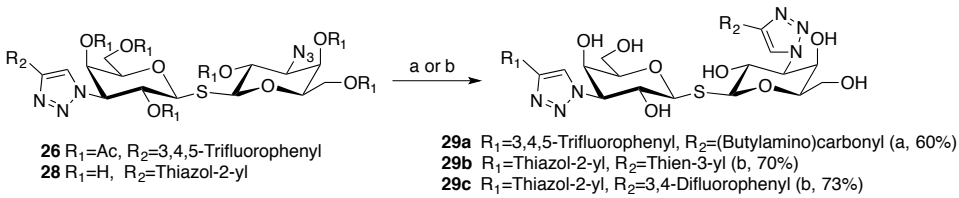
3.3.3 Synthesis of high affinity thiodigalactosides

Based on the inhibition potencies (Chapter 4) of triazoles **17a-r** and **20a-d** towards galectin-1 and galectin-3, a selection of symmetric bis-3-(4-aryl-1,2,3-triazol-1-yl)-thiodigalactosides **24a-e** were synthesized from diazide **23**[56] through Cu(I)-catalyzed Huisgen 1,3-dipolar cycloadditions (Scheme 5). Using only one alkyne equivalent resulted in mono-3-(4-aryl-1,2,3-triazol-1-yl)-thiodigalactosides **25-28**, later used to make asymmetric thiodigalactosides. However, an approximate 1:2:1 product distribution of starting material, mono-cycloadduct and di-cycloadduct can explain the low yields of **25**, **26** and **28**. The strategy of using only one alkyne equivalent is not material efficient, but it is a quick route to the desired intermediates.



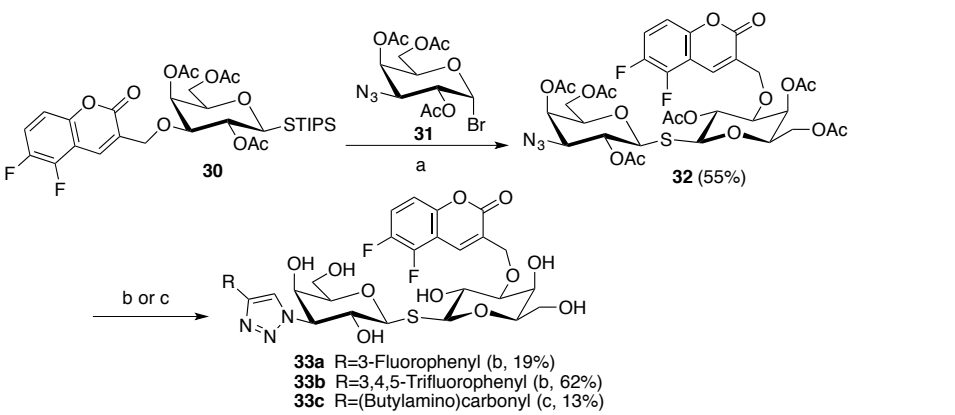
Scheme 5. Synthesis of thiodigalactosides **24a-e** and **25-28**. Reagents and conditions: a) i. alkyne, CuI, DIPEA, DMF, 50 °C; ii. NaOMe, MeOH, DCM, rt; b) alkyne, CuI, DIPEA, DMF, 50 °C; c) NaOMe, MeOH, DCM, rt.

The crystal structures of the galectin-3:**7** and galectin-1:**8** complexes both revealed favorable interactions with each aryltriazolyl group, making symmetrical thiodigalactosides suitable as galectin ligands. Since the synthesis of symmetrical thiodigalactosides is easy and quick, why then do the more time consuming synthesis of asymmetrical thiodigalactosides? One reason is that synthesis of asymmetrical thiodigalactosides enables the combination of different aryls *e.g.* a trifluorophenyltriazole (high galectin-3 affinity) with a difluorocoumaryl group[32] (high galectin-3 selectivity) or with a butylamidotriazole[25] (high galectin-3 affinity) or with a fluorescent tag. Since any changes made to a symmetrical thiodigalactoside will affect the binding interactions in both subsites A-B and D-E, conclusions about either side of the binding pocket is difficult to draw. Another reason for making asymmetrical thiodigalactosides is that it gives information of the binding interactions in subsites D-E. Asymmetric bis-3-(4-aryl-1,2,3-triazol-1-yl)-thiodigalactosides **29a-c** were synthesized through a 1,3-dipolar cycloaddition with either thiodigalactoside **26** or **28** (Scheme 6) followed by deacetylation with butylamine in case of **29a**.



Scheme 6. Synthesis of thiodigalactosides **29a-c**. Reagents and conditions: a) From **26**: i. methyl propiolate, CuI, DIPEA, MeCN, 50 °C; ii. $BuNH_2$, MeOH, rt; b) From **28**: alkyne, CuI, DIPEA, DMF, 50 °C.

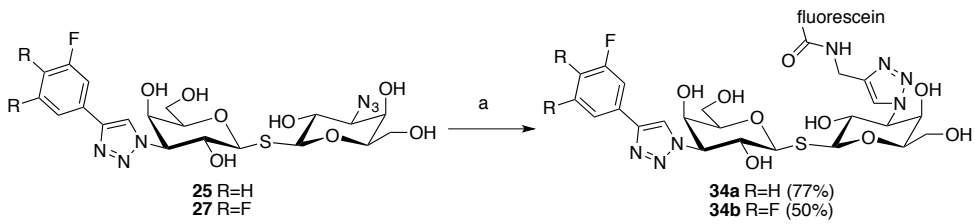
Introducing both a triazole and a coumaryl group at a thiodigalactoside derivative is more difficult since both reactions involve an alkyne, an azide and a copper catalyst. An alternative approach to synthesize asymmetric thiodigalactosides is to introduce one, or both, C3-substituents at the monosaccharide level followed by *S*-glycosylation, which is how the combination of a difluorocoumaryl group with different phenyltriazoles was achieved. *S*-glycosylation of glycosyl donor **30**[55] with glycosyl acceptor **31**[54] resulted in thiodigalactoside **32** and reaction with three different alkynes and subsequent deprotection (Scheme 7) gave asymmetric thiodigalactosides **33a-c**.



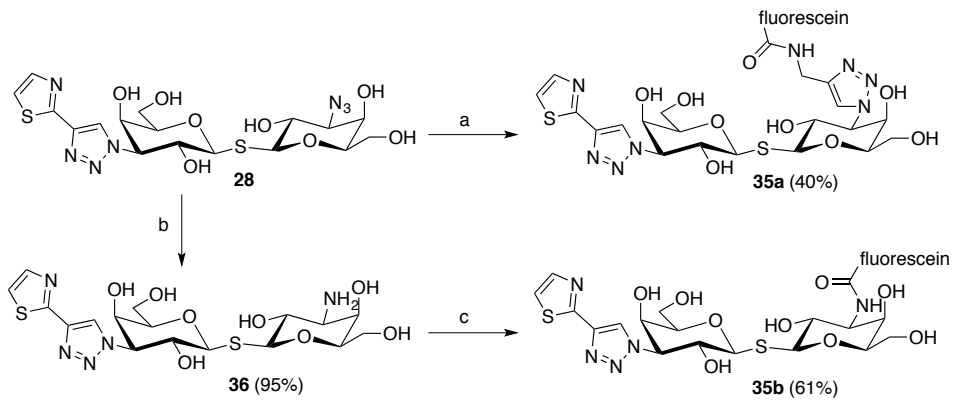
Scheme 7. Synthesis of thiodigalactosides **33a-c**. Reagents and conditions: a) TBAF (1M in THF), MeCN, rt; b) i. alkyne, CuI, Et_3N or DIPEA, MeCN or DMF, 50 °C; ii. NaOMe, MeOH, DCM, rt; c) i. methyl propiolate, CuI, DIPEA, MeCN, 50 °C; ii. $BuNH_2$, MeOH, rt.

3.4 Synthesis of fluorescent probes (Papers II-III)

A key component in the competitive fluorescence polarization assay (Chapter 4) used to evaluate the affinity of synthesized galectin ligands is the fluorescent probe that the ligand competes with. In order to more accurately evaluate high affinity ligands, the fluorescent probes also need to have high affinity. Hence, we decided to synthesize fluorescent probes **34a-b** (Scheme 8, towards galectin-3) and **35a** (Scheme 9, towards galectin-1) with a fluorescein instead of an aryl at one of the triazoles, introduced by a 1,3-dipolar cycloaddition from either azide **25**, **27** or **28**. If the fluorescent moiety is too flexible it will, due to propeller effects (see section 4.1), have a negative impact on the fluorescence polarization assay, which was the case of probe **35a**. We hypothesized that removing the triazole-linker in **35a** would reduce the flexibility of the fluorescein, thus probe **35b** (Scheme 9) was synthesized through reduction of azide **28** to amine **36** followed by an amide coupling.



Scheme 8. Synthesis of probes **34a-b**. Reagents and conditions: a) 5-FAM alkyne, CuI, DIPEA, DMF, 50 °C.



Scheme 9. Synthesis of probes **35a-b**. Reagents and conditions: a) 5-FAM alkyne, CuI, DIPEA, DMF, 50 °C; b) 1,3-propanedithiol, Et_3N, MeOH, rt; c) 5-FAM-NHS, DIPEA, DMSO, rt.

4. Evaluation of galectin ligand affinity and selectivity

The first aim of this thesis was to synthesize galectin ligands with high affinity. So how do we measure affinity? There are several methods available[41] with different advantages and disadvantages. We have used competitive fluorescence polarization (FP)[55,57] to measure galectin affinities and the most potent galectin-3 ligands were also measured with competitive isothermal titration calorimetry (ITC)[55].

4.1 Competitive fluorescence polarization

One of the advantages of measuring molecular binding with FP is that it directly measures the bound/free ratio of the labeled molecule, while most other methods require separation of the bound and free labeled molecule before measuring. In FP measurements a fluorescently labeled molecule is excited with plane-polarized light, that, if the molecule remains stationary, emit back into a fixed plane. However, since molecules can rotate and tumble, some of the light will be emitted into a different plane, which is referred to as depolarization. The only parameter that affects molecular motion, if temperature and viscosity are kept constant, is molecular volume. Smaller molecules rotate and tumble faster than larger ones. This means that an excited small fluorescently labeled molecule will be more depolarized when free in solution than when bound to a larger protein (Figure 11).[58] The polarization, or anisotropy, is measured and the amount of bound and free state of a fluorescently labeled molecule is calculated[57], and used to determine the dissociation constant (K_d) of the fluorescently labeled molecule.

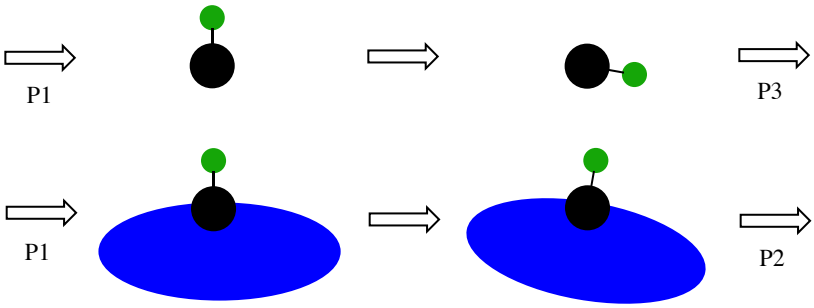


Figure 11. Mechanistic overview of fluorescence polarization. The fluorescently labeled molecules free in solution (upper) and bound to a protein (lower) are first excited with polarized light (P1) followed by size-dependent tumbling and finally emission of polarized light (P2 and P3) where P1>P2>P3.

Fluorescent labeling of every ligand would be both impractical and expensive and to avoid this, FP measurements of non-labeled ligands are performed in a competitive manner together with a fluorescently labeled probe. By running the FP assay with known concentrations of ligand, galectin and probe (with a known K_d) the K_d of the non-labeled ligand can be calculated[57]. This approach allows for fast, inexpensive and reliable affinity determination of synthesized ligands.

The probe needs to bind the galectin in the same location as the ligand it competes with, and it should preferably bind equally or better than the ligand being evaluated. The probe should also give a high A_{max} value, which is the measured anisotropy when the galectin is saturated with the probe. This is to create a good span between the values of bound and unbound probe, which makes calculations from anisotropy values less sensitive to errors of measurement. However, the A_{max} will be lowered if the fluorescent moiety tumbles faster than the rest of the bound probe molecule, which is known as propeller effects.[57] If the ligand binds much stronger than the probe, then the titration curve goes from sigmoidal to more stepwise and thus increasing the errors in the K_d calculations. The probes **34a-b** were evaluated towards galectin-3 in direct FP titration (Figure 12A-B) and found to have satisfactory K_d and A_{max}. Probe **34b** was also evaluated towards galectin-3C (Figure 12C). Probe **34b** was used to accurately evaluate galectin-3 ligands with K_d in the low nM range. The use of fluorescent probe with high affinity also allows lower galectin and ligand concentrations to be used in the FP measurements. The lower galectin concentration requires the use of a blocking protein, in our case BSA, that coats the microwells to prevent perturbing losses of galectin due to absorption to microwells.

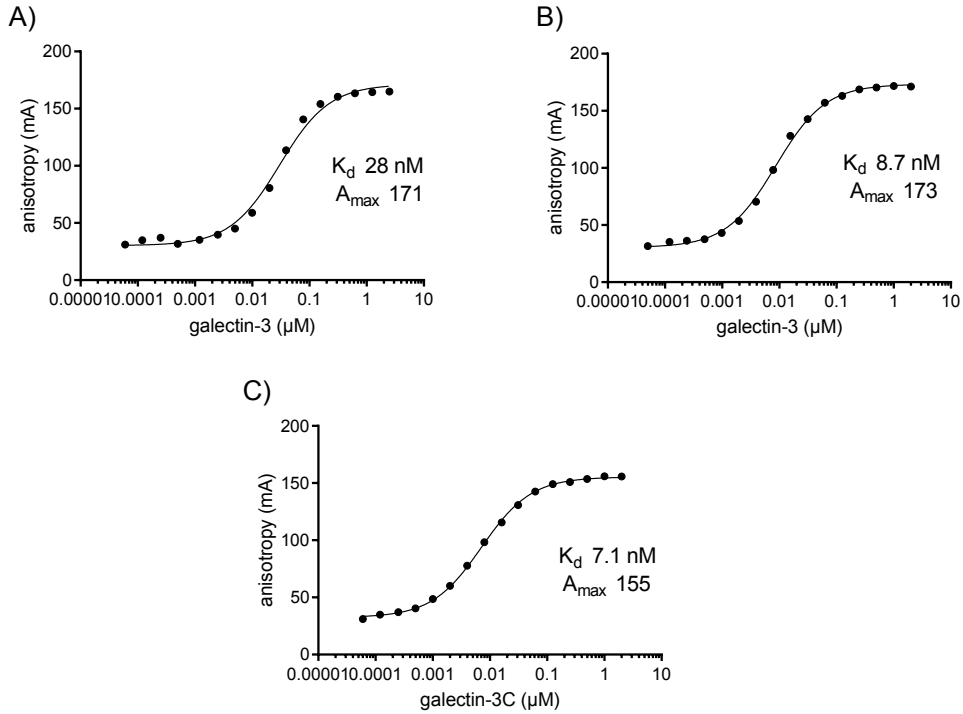


Figure 12. Fluorescence polarization titration of galectin-3 with the fluorescent probes A) **34a** and B) **34b**. C) Fluorescence polarization titration of galectin-3C with the fluorescent probe **34b**.

Probe **35a**, analogous to **34a-b**, was evaluated towards galectin-1 but its A_{max} value (Figure 13A) was deemed too low to be used in competitive FP. In order to increase the A_{max} value, the flexibility of the fluorescein moiety was reduced through removal of the triazole-linker in **35a** resulting in probe **35b** (Figure 13B). Probe **35b** did indeed have a higher A_{max} value allowing lower galectin-1 and ligand concentrations in evaluation of galectin-1 ligands with low nM affinity.

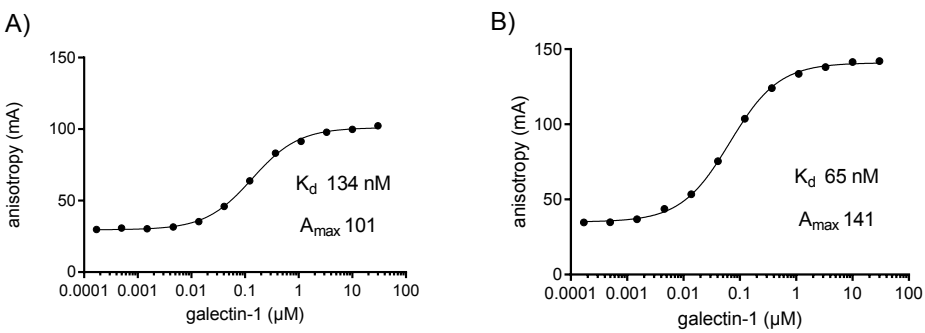


Figure 13. Fluorescence polarization titration of galectin-1 with the fluorescent probes A) **35a** and B) **35b**.

4.2 Competitive isothermal titration calorimetry

In ITC the heat released (exothermic) or absorbed (endothermic) when a ligand binds a protein is measured using a reference cell and a sample cell connected by a circuit. Both cells have heaters to keep the temperatures identical and when something is injected into the sample cell that results in a temperature change the heater adjusts its input of power to maintain identical temperature with the reference cell. This input of power is measured in an ITC experiment resulting in a thermogram and a titration curve.[59] In a typical ITC experiment the protein is in the sample cell and the ligand is being titrated in, but the reverse is also possible. Great care must be taken when preparing the protein- and ligand solutions since small deviations in buffer or DMSO content between the sample cell and the titration solution will give rise to a heat signal typically bigger than the binding event. ITC requires more protein than FP, as evident by the thiodigalactosides that were analyzed with both ITC and FP, where the galectin-3 concentration in ITC was 11 μM but only 10 nM in FP. However, ITC has advantages such as also directly measuring the bound/free ratio of the ligand without separation and it does not require any labeling of the ligands. But more importantly, while FP only gives the K_d, ITC additionally provides the stoichiometry and the enthalpy (ΔH), from which the entropy (ΔS) can be calculated. If measurements are repeated at different temperatures the heat capacity (ΔC_p) can also be obtained.[60] The additional information about the enthalpy- and entropy contributions is very useful when elucidating how a ligand binds to a protein. The ideal curve for an ITC experiment is sigmoidal; first all the titrated ligand binds to the protein giving the lower plateau then as more ligand is added to the protein an equilibrium sets in and the protein starts to be saturated giving the slope in the curve and finally when the protein is saturated the curve flattens out. The enthalpy is obtained from the lower plateau and the K_d is obtained from the slope of the curve.

Lets look at an example where thiodigalactoside **7** (K_d = 2.3 nM) is titrated into galectin-3C (Figure 14A). What is observed here is not a sigmoidal curve but more of a stepwise curve, this leads to bigger errors in estimation of the K_d and entropy that are obtained from the slope. A solution to this is to do competitive ITC where the protein is saturated with a weaker ligand with known thermodynamic parameters before titrating in a ligand of unknown affinity. We measured thiodigalactoside **7** again, but this time the galectin-3C had been saturated with galactosylthioglucoside **22a** (K_d = 1900 nM) and this time a sigmoidal curve was obtained (Figure 14B). This competitive ITC protocol[55] was used to determine the K_d of our most potent galectin-3 ligands and as galectin-3C was used instead of galectin-3, to minimize oligomerization, the ligands were also evaluated towards galectin-3C using FP and probe molecule **34b**.

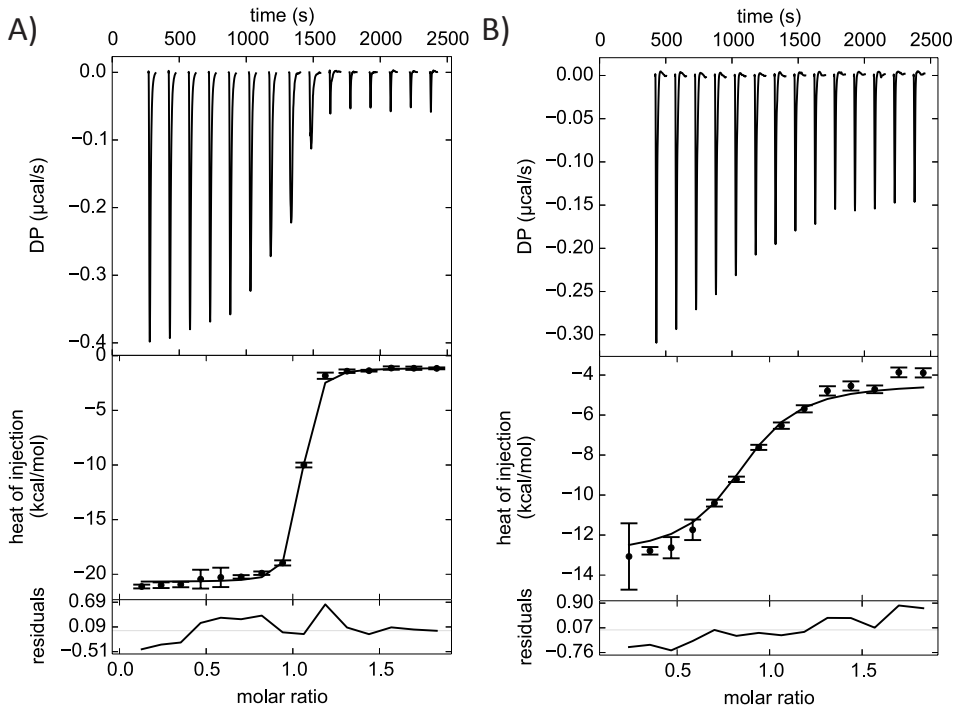


Figure 14. A) Direct ITC with **7** and galectin-3C revealing a step titration slope indicative of high affinity and providing lower accuracy extraction of binding parameters. B) Competitive ITC with **7** and galectin-3C saturated with **22a** revealing a sigmoidal titration curve allowing for accurate extraction of binding parameters.

Figure 14 shows the benefits of using competitive ITC when evaluating high affinity ligands, however, there is a downside. The use of an additional component increases the sources of errors. Because solutions of every component need to be prepared and exact concentrations are not always achieved.

4.3 Binding data (Papers II-IV)

Based on the galectin affinities of previously synthesized bis-3-(4-aryl-1,2,3-triazol-1-yl)-thiodigalactosides[27], ligand evaluation was limited to galectin-1 and galectin-3. The inhibition potencies of the synthesized ligands were determined by competitive FP[57]. In the competitive FP assay, the galectin ligands were dissolved in DMSO at 10-50 mM and diluted in PBS to 3-6 different concentrations (<5 % DMSO) and tested in duplicates, K_d and standard error of the mean (SEM) were calculated from 4 to 25 single point measurements from at least two independent

experiments showing between 10-90% inhibition. The SEM was calculated using the following equations:

$$SEM = \frac{\sigma}{\sqrt{n}}$$

where *n* is the number of data points and σ is the standard deviation:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x_{mean})^2}{n-1}}$$

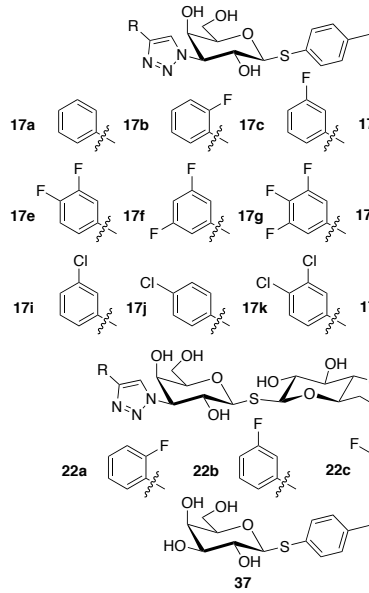
where x_i is the individual data points and x_{mean} is the mean of all data points.

The most potent galectin-3 ligands, thiodigalactosides **24a-c**, **29a** and **33a-b**, were also evaluated with a second binding assay, competitive ITC[55], to corroborate their binding affinities and to determine binding thermodynamic parameters towards galectin-3.

4.3.1 Monosaccharides

To study the binding interactions in subsite A of galectin-1 and galectin-3 thiogalactosides **17a-r**, **20a-d** and **22a-c** were evaluated together with reference compounds *p*-methylphenyl 1-thio-β-D-galactopyranoside **37** and carboxymethyl 1-thio-β-D-galactopyranoside **38**[61] (Table 1 and 2). All phenyltriazoles **17a-l** bound galectin-3 with higher affinity than reference **37**. Increased galectin-3 affinity, compared to the unsubstituted phenyl **17a**, was observed with fluorination (**17c-g**), chlorination (**17i-k**) and bromination (**17l**) in the meta and/or para positions of the phenyl ring, while ortho-fluorination (**17b**) did not increase affinity compared to **17a**. The increase in galectin-3 affinity compared to **17a** was 3-4 fold for monofluorination (**17c-d** and **17h**), 6-17 fold for di- or trifluorination (**17e-g**), 2-6 fold for mono- or dichlorination (**17i-k**) and 2-fold for bromination (**17l**). The highest affinity (5.2 μM) was observed for the trifluorophenyltriazole **17g** and multiple fluorinations on the phenyl (**17e-g**) increased the galectin-3 selectivity, while multiple chlorinations (**17k**) did not. The galectin-3 affinity of galactosylthioglucosides **22a-c** was higher than the corresponding *p*-methylphenyl 1-thio-β-D-galactopyranosides, but the affinity contribution from the mono-fluorinated phenyltriazole moieties had a similar meta>para>ortho trend (1:1.3:4.1), indicating similar binding interactions.

Table 1. K_d (μM) values and SEM for thiogalactosides **17a-l** and **22a-c** and reference compound **37**.

	Galectin-1	Galectin-3	Galectin-1/ Galectin-3	Structures
37	1110 ± 110	230 ± 30	4.7	
17a	220 ± 6	88 ± 3	2.5	
17b	190 ± 13	92 ± 5	2.1	
17c	180 ± 7	22 ± 0.7	8.2	
17d	140 ± 6	31 ± 1.3	4.5	
17e	530 ± 40	8.8 ± 0.3	60	
17f	290 ± 9	15 ± 0.3	19	
17g	180 ± 12	5.2 ± 0.3	35	
17h	170 ± 2	23 ± 1.0	7.4	
17i	100 ± 2	44 ± 1.1	2.3	
17j	680 ± 80	19 ± 1.0	36	
17k	120 ± 8	14 ± 1.3	8.6	
17l	300 ± 20	44 ± 3.9	6.8	
22a	nd	1.9 ± 0.1	-	
22b	nd	0.46 ± 0.02	-	
22c	nd	0.60 ± 0.02	-	

nd = not determined

Replacing the phenyl in **17a-l** with different five-membered heterocycles resulted in thiogalactosides **17m-r** and **20a-d** that showed higher galectin-1 affinity than the corresponding reference (**37** or **38**). The higher galectin-1 affinity of five-membered heterocycles compared to substituted phenyls is due to the smaller binding pocket in subsite A of galectin-1 (Figure 15). The aliphatic cyclopentyl (**17m**) bound 2-8 fold worse than the aryltriazoles (**17n-r**) suggesting that aliphatic ring systems are less favored than aromatic ring systems. The imidazol-2-yltriazole (**17q**) and thiazol-4-yltriazole (**20c**) had galectin-1 affinity equal to their corresponding thien-3-yltriazole (**17p** or **20a**), while the thiazol-2-yltriazole (**20b**) exhibited the highest galectin-1 affinity. The binding data also revealed that an oxygen (**17n**) or a nitrogen (**17r** and **20d**) placed in the 3-position from the triazole reduced the galectin-1 affinity 2-3 fold compared to the corresponding thien-3-yl analog. Thien-3-yltriazole **17p** bound galectin-1 slightly better than thien-3-yltriazole **20a**, but this does not correlate with references **37** and **38**, indicating an interplay between substituents when bound to their respective subsite A-B and D.

Table 2. K_d (μM) values and SEM for thiogalactosides **17m-r** and **20a-d** and reference compounds **37-38**.

	Galectin-1	Galectin-3	Galectin-3/ Galectin-1	Structures
37	1110 ± 110	230 ± 30	0.2	
17m	370 ± 30	980 ± 60	2.6	
17n	120 ± 7	250 ± 20	2.1	
17o	61 ± 4	180 ± 10	3.0	
17p	47 ± 2	200 ± 10	4.3	
17q	47 ± 3	220 ± 10	4.7	
17r	140 ± 7	330 ± 20	2.4	
38	430 ± 40	1800 ± 100	4.2	
20a	58 ± 4	200 ± 7	3.4	
20b	43 ± 2	280 ± 7	4.5	
20c	62 ± 2	540 ± 30	8.7	
20d	180 ± 9	430 ± 20	2.4	

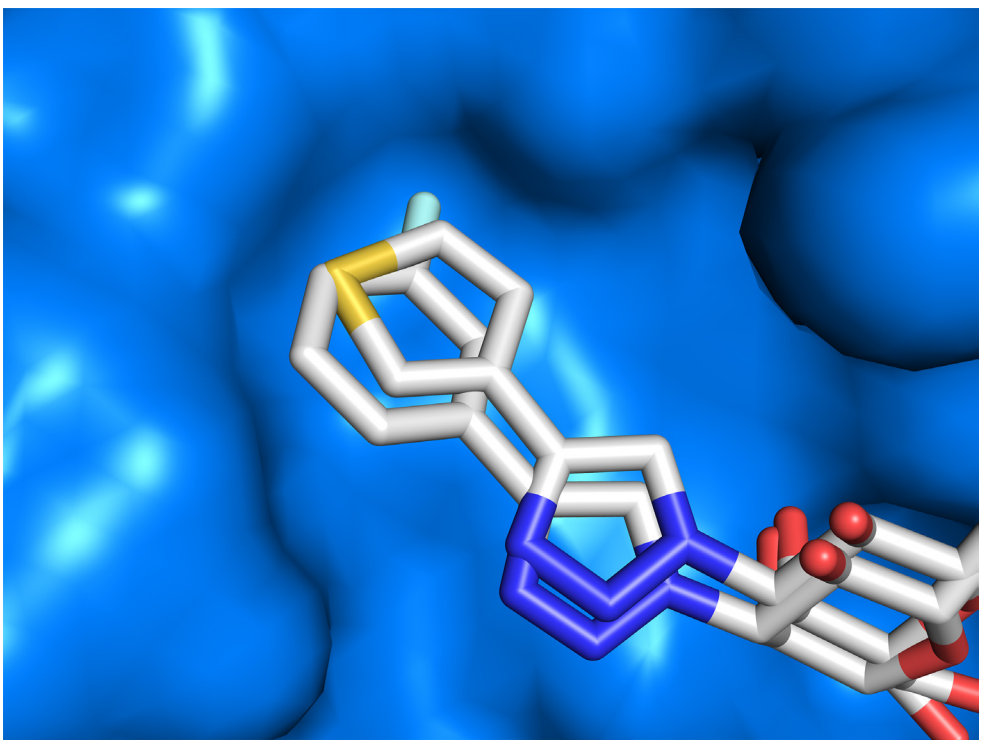
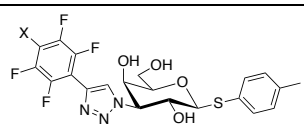


Figure 15. Close-up view of subsite A in the crystal structure of galectin-1 in complex with thiodigalactosides **7**[62]-**8**[Paper III]. Illustrating the smaller binding pocket in subsite A of galectin-1.

Evaluation of penta- and tetrafluorophenyltriazoles **18a-i** towards galectin-3 with FP (Table 3) revealed a 3.4 μM galectin-3 affinity for pentafluorophenyl **18b**, while all substitutions of the fluorine in para position led to a drop in affinity. This is not surprising since the fluorine in para position is involved in an orthogonal multipolar fluorine-amide interaction with Ser237.[27] Azide **18c** and amine **18d** resulted in a 2-3 fold decrease in affinity, while replacement with a hydroxyl (**18a**) surprisingly resulted in a 7-fold decrease. The addition of methyl groups (**18g-h**) to amine **18d** further decreased the affinity 2-fold per added methyl. Replacement of fluorine with a methoxy group (**18e**) gave similar affinity as hydroxyl **18a**, while replacement with a bulkier ethoxy group (**18f**) resulted in almost 5-fold lower affinity compared to the methoxy **18e**, which is indicative of steric restrictions in the binding pocket. This would also explain why the even bulkier pyrrolidine **18i** does not bind galectin-3 at the highest measured concentration (precipitation of the ligand was also observed at this concentration).

Table 3. K_d (μM) values and SEM for thiogalactosides **18a-i**.

	Galectin-3	Structures
18a	23 ± 1.7	
18b	3.4 ± 0.21	
18c	8.5 ± 1.2	
18d	11 ± 0.6	
18e	18 ± 2.1	
18f	88 ± 12	
18g	18 ± 0.9	
18h	40 ± 3.3	
18i	>300	

18a X=OH
18b X=F
18c X=N_3
18d X=NH_2
18e X=OMe
18f X=OEt
18g X=NHMe
18h X=NMe_2
18i X=pyrrolidine

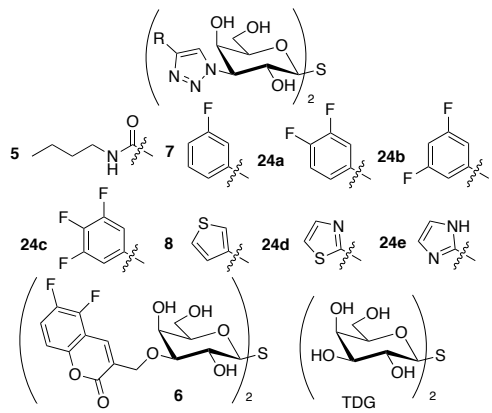
4.3.2 Thiodigalactosides

Evaluation of symmetrical thiodigalactosides **24a-e** together with reference thiodigalactosides **5-8**[25,27,32], **39**[27] and TDG[24] with FP (Table 4) revealed single-digit nM affinities towards galectin-3 (**24a-c**) and galectin-1 (**24d**). The di- and trifluorophenyltriazoles **24b-c** exhibited 30-40 fold galectin-3 selectivity while the 3,4-difluoropheyltriazole **24a** proved to be unselective, which is surprising considering the monosaccharide (**17e**) that had a 60-fold galectin-3 selectivity. The loss of galectin-3 selectivity for **24a** indicates a favorable interaction in subsite E (Figures 9-10) of galectin-1, which provided the reason for making asymmetrical thiodigalactoside **29c** (combination of a thiazol-2-yltriazole and a 3,4-difluorophenyltriazole). Interestingly, the imidazol-2-yltriazole **24e** showed 4-fold lower affinity than the thien-3-yltriazole **8**, while they had equal affinity as

monosaccharides (*c.f.* **17p** and **17q** in table 2). This indicates a difference in binding interactions in subsite E. The observed 7-10 fold galectin-1 selectivity of thiodigalactosides **8** and **24d-e** is twice that of monosaccharides **17p-q** and **20a-b** indicating a preference for five-membered heterocycles not only in subsite A, but also in subsite E of galectin-1 compared to galectin-3.

Table 4. K_d (nM) values and SEM for symmetrical thiodigalactosides TDG, **5-8**, **24a-e** and **39** determined by competitive fluorescence polarization.

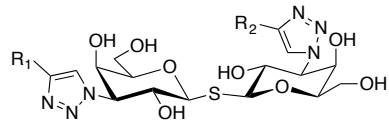
	Galectin-1	Galectin-3	Galectin-1/ Galectin-3	Structures
TDG	24 000[24]	49 000[24]	0.5	
5	230 ± 30	99 ± 4	2.3	
6	3 900[32]	57 ± 4	68	
7	12[27]	2.3 ± 0.2	5.2	
39	27[27]	4.0 ± 0.6	6.8	
24a	<10	1.1 ± 0.2	<9	
24b	63 ± 15	1.6 ± 0.3	39	
24c	69 ± 5	2.3 ± 0.4	30	
8	6.1 ± 1	59 ± 4	0.10	
24d	8.4 ± 1	74 ± 4	0.11	
24e	25 ± 2	179 ± 13	0.14	



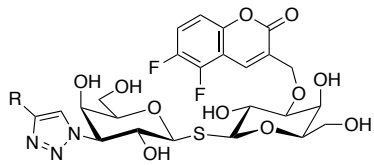
The combination of different aryl substituents at the C3-positions of TDG resulting in asymmetrical thiodigalactosides **29a-c** and **33a-c**, allowed fine-tuning of galectin affinities. Evaluation with FP (Table 5) revealed single-digit nM galectin-3 affinities (**29a**, **29c**, **33b** and **34b**) and low nM galectin-1 affinities (**29b-c**). Asymmetric bis-3-(4-aryl-1,2,3-triazol-1-yl)-thiodigalactoside **29a** showed similar binding affinity and selectivity as symmetric bis-3-(4-aryl-1,2,3-triazol-1-yl)-thiodigalactosides **24b-c**. The asymmetric thiodigalactoside **33b** (with a coumaryl group) had improved galectin-3 selectivity, while maintaining single-digit nM affinity. The fluorescent probes (**34a-b**, Scheme 8) further demonstrated the galectin-3 affinity enhancement of additional fluorination on the phenyl ring as trifluorinated probe **34b** had 3-fold higher galectin-3 affinity than monofluorinated probe **34a**. Probe **34b** also displayed a remarkable 57-fold galectin-3 selectivity. The combination of a thiazol-2-yltriazole and a thien-3-yltriazole (**29b**) resulted in 12 nM galectin-1 affinity, while the combination of a 2-thiazole and a 3,4-difluorophenyl (**29c**) resulted in higher affinity towards galectin-3 than galectin-1.

Table 5. K_d (nM) values and SEM for asymmetrical thiodigalactosides **29a-c** and **33a-c** determined by competitive fluorescence polarization, with symmetrical thiodigalactosides **5-8** for comparison, and for probes **34a-b** with direct fluorescence polarization titration.

	Galectin-1	Galectin-3	Galectin-1/ Galectin-3	Structures
5	230 ± 30	99 ± 4	2.3	
6	3 900[32]	57 ± 4	68	
7	12[27]	2.3 ± 0.2	5.2	
8	6.1 ± 1	59 ± 4	0.10	
29a	63 ± 7	2.4 ± 0.2	26	
29b	12 ± 1	61 ± 3	0.20	
29c	13 ± 1	1.8 ± 0.2	7.2	
33a	260 ± 20	12 ± 1	21	
33b	340 ± 20	7.5 ± 0.7	46	
33c	660 ± 50	47 ± 3	14	
34a	390 ± 12	28 ± 2	14	
34b	500 ± 15	8.7 ± 0.4	57	



29a R_1=3,4,5-Trifluorophenyl, R_2=(Butylamino)carbonyl
29b R_1=Thiazol-2-yl, R_2=Thien-3-yl
29c R_1=Thiazol-2-yl, R_2=3,4-Difluorophenyl



33a R=3-Fluorophenyl
33b R=3,4,5-Trifluorophenyl
33c R=(Butylamino)carbonyl

Differentiating between highly potent ligands can be difficult in any binding assay, since very low concentrations are used and measurements must be precise. Hence, thiodigalactosides **7**, **24a-c**, **29a**, **33b** and **39**, all with single-digit nM affinity towards galectin-3, were evaluated with a second independent binding assay, namely competitive ITC (Table 6, Figure 16). The galectin-3C K_d values obtained from FP matched those for galectin-3 (*c.f.* Tables 4 and 5), demonstrating that the ligand binding properties of galectin-3C are essentially identical to those of galectin-3. The ITC measurements gave low nM affinities for all thiodigalactosides, but some (**7**, **29a** and **33b**) K_d values were higher than what was obtained with FP. Based on these measurements thiodigalactosides **24a-c** were determined as more potent galectin-3 ligands than the other thiodigalactosides, but we are still not able to differentiate between thiodigalactosides **24a-c**. The ITC measurements also reveal that the binding for all thiodigalactosides is driven by a large enthalpic contribution, while the entropy is unfavorable upon binding. Although within the margin of error, it seems that the 4-fluorophenyl **39** has a larger enthalpic contribution than the 3-fluorophenyl **7**, which is counteracted by a more unfavorable entropic term. The 8.4 kJ/mol larger enthalpic contribution of 3,4-difluorophenyl **24a** compared to 3,5-difluorophenyl **24b**, however, is explained by the additional fluorine-amide interaction of the para fluorine in **24a**, while the second meta fluorine in **24b** points towards solution (Figure 19 in section 5.1).

Table 6. Evaluation of thiodigalactosides **7**, **24a-c**, **29a**, **33b** and **39** with FP and competitive ITC towards galectin-3C. Errors in parentheses showing a 68% confidence interval.

	K_d FP (nM)	K_d ITC (nM)	ΔG (kJ/mol)	ΔH (kJ/mol)	-TΔS (kJ/mol)
7	2.3 ± 0.2	8.7 (5.2, 15)	-46.5 (-47.5, -44.8)	-73.1 (-78.6, -69.4)	26.6 (21.0, 30.7)
39	2.7 ± 0.3	21 (14, 35)	-44.3 (-45.2, -42.6)	-77.3 (-86.3, -72.4)	33.1 (24.1, 38.3)
24a	<1	2.9 (1.9, 4.2)	-49.3 (-50.1, -48.2)	-79.7 (-83.3, -76.7)	30.4 (26.7, 33.6)
24b	<1	3.2 (1.6, 6.1)	-49.0 (-50.3, -46.7)	-71.3 (-76.0, -68.0)	22.3 (17.4, 26.4)
24c	<1	1.7 (1.0, 2.7)	-50.6 (-51.6, -49.1)	-74.8 (-78.1, -72.2)	24.2 (20.8, 27.2)
29a	1.5 ± 0.4	19 (13, 31)	-44.6 (-45.4, -43.0)	-75.2 (-81.9, -71.6)	30.6 (23.8, 34.5)
33b	4.0 ± 0.4	11 (0.14, 19)	-45.9 (-48.4, -44.1)	-73.8 (-78.2, -68.2)	27.9 (22.9, 33.8)

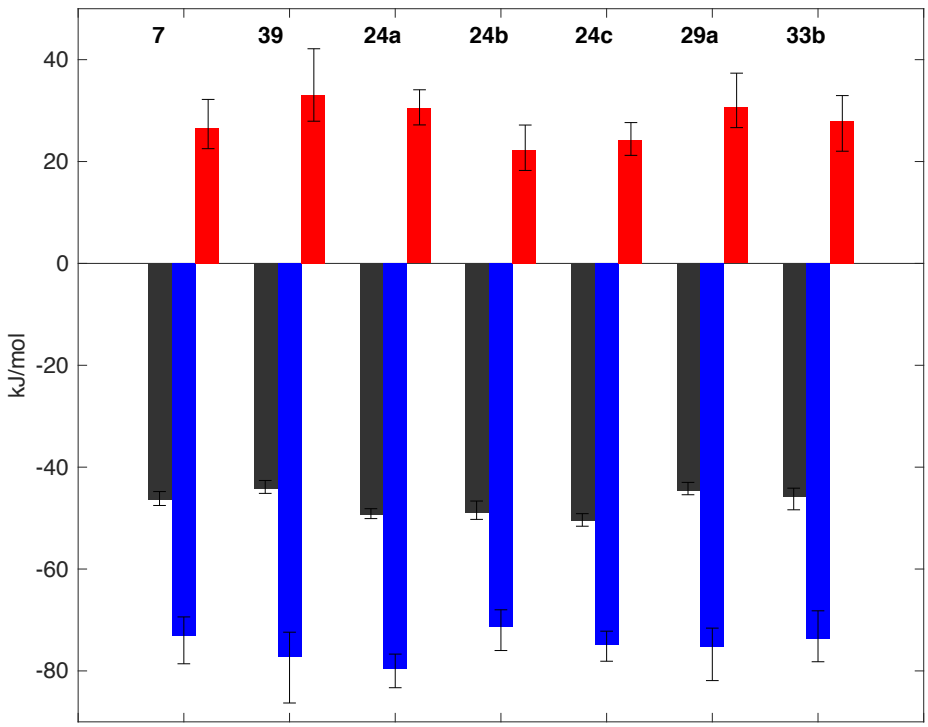


Figure 16. Thermodynamic parameters, ΔG (black bars), ΔH (blue bars) and -TΔS (red bars), of **7**, **24a-c**, **33b** and **39** as determined in competitive ITC experiments with **22a**.

To summarize, polyfluorinated phenyltriazoles favor galectin-3 binding, while thien-3-yltriazole and thiazol-2-yltriazole favor galectin-1 binding (Figure 17). The 3,4,5-trifluorophenyltriazole also showed improved galectin-3 selectivity, especially in combination with the difluorocoumaryl group. Differentiation between the most potent galectin-3 ligands was not possible even by using two independent binding assays. However, we were able to differentiate between several of the single-digit nM galectin-3 ligands.

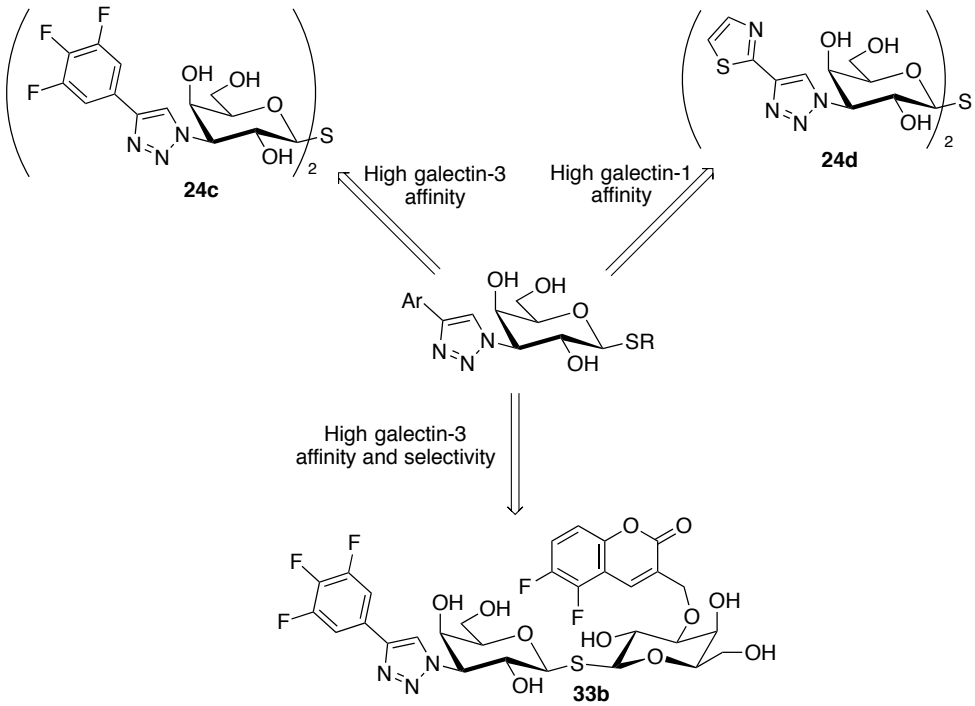


Figure 17. Summary of the optimization of the aryl groups at the 3- and 3'- positions of thiodigalactoside.

5. Deciphering the binding of galectin-3 ligands

The binding interactions of TDG to galectin-3 has previously[63] been determined with X-ray crystallography (Figure 18). The pyranose ring of the galactose that occupies subsite C, forms CH-π stacking interactions with Trp181, and the hydroxyl groups at C4 and C6 are involved in hydrogen bonding with His158, Asn160, Arg162, Asn174 and Glu184. The galactose in subsite D has fewer interactions since much of the ring (C4-C6) is solvent exposed with only the hydroxyl group at C2 being involved in hydrogen bonding with Arg162, Glu184 and Arg186.

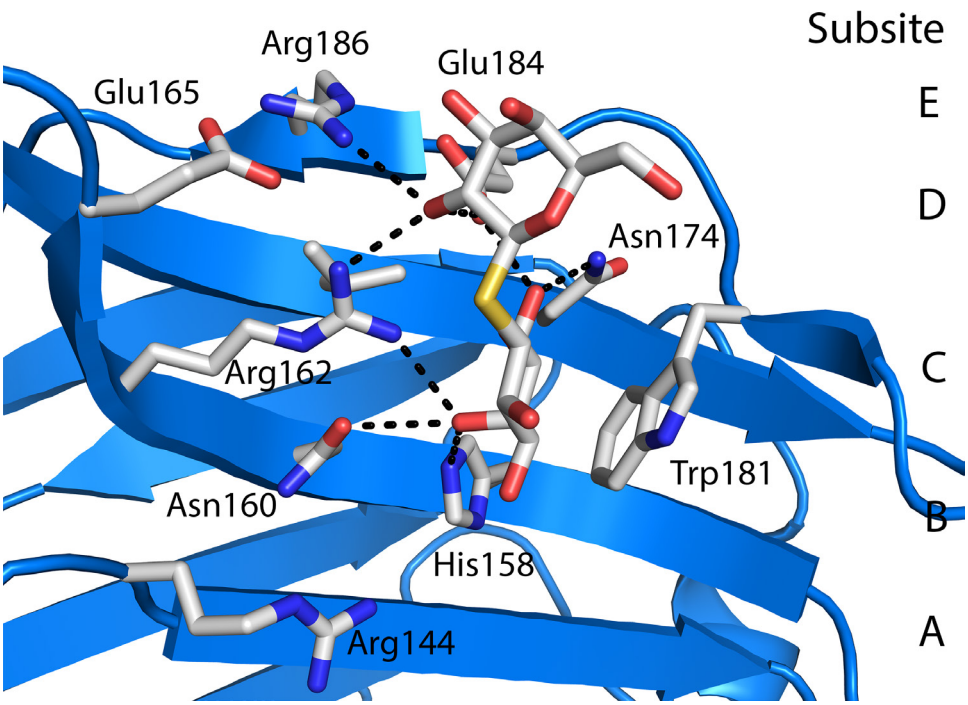


Figure 18. Crystal structure of TDG in complex with galectin-3C. Hydrogen bonds are indicated with black dashed lines.

5.1 Fluorine-amide and arginine-π interactions (Paper V)

Crystal structures of thiodigalactosides **7** and **39** in complex with galectin-3 have revealed[27] orthogonal multipolar fluorine-amide interactions with backbone amides of Arg144, Ile145, and Ser237 and cation-π interactions with Arg144 and Arg186. We wanted to look more closely at these interactions and determine the importance of each of them using X-ray crystallography, ITC, and mutant studies with FP.

Looking back at phenyltriazoles **17a-d**, we observed that the addition of a fluorine in the ortho position (**17b**) did not improve galectin-3 affinity compared to the unsubstituted phenyl (**17a**), while meta- (**17c**) and para (**17d**) fluorination increased the galectin-3 affinity 4-fold and 3-fold, respectively, compared to **17a**. This improvement was also observed for galactosylthioglucosides **22a-c**. To explain this, the crystal structures (Figure 19) of phenyltriazoles **17b-d** in complex with galectin-3C were determined. The fluorine in the galectin-3:**17b** complex is positioned almost orthogonal to the side-chain amide of Asn160 at a 3.5 Å distance, which suggest a possible fluorine-amide interaction. The galectin-3:**17c** and galectin-3:**17d** complexes revealed the previously reported[27] orthogonal multipolar fluorine-amide interactions with backbone amides Arg144 and Ile145 for **17c** and Ser237 for **17d**. The galectin-3:**17d** complex also revealed a favorable multipolar C-F···H-C_α interaction with the hydrogen at the α-carbon of Gly238, where the distance between the fluorine and the α-carbon of Gly238 was 2.9 Å. When the structures were superimposed (Figure 19D) we observed that the phenyl ring in **17d** had shifted slightly compared to the other structures. This shift positions the ligand-fluorine between the Ser237 and Gly238, which may optimize the interactions of **17d** with Ser237 and Gly238.

The observed orthogonal multipolar fluorine-amide (backbone) interactions and the multipolar C-F···H-C_α interaction may explain the 3-4 fold increase in galectin-3 affinity of **17c-d** compared to the unsubstituted phenyl **17a** (Table 1). However, ortho fluorination (**17b**) did not increase galectin-3 affinity, suggesting a lack of a favorable fluorine-amide interaction with the side-chain amide of Asn160.

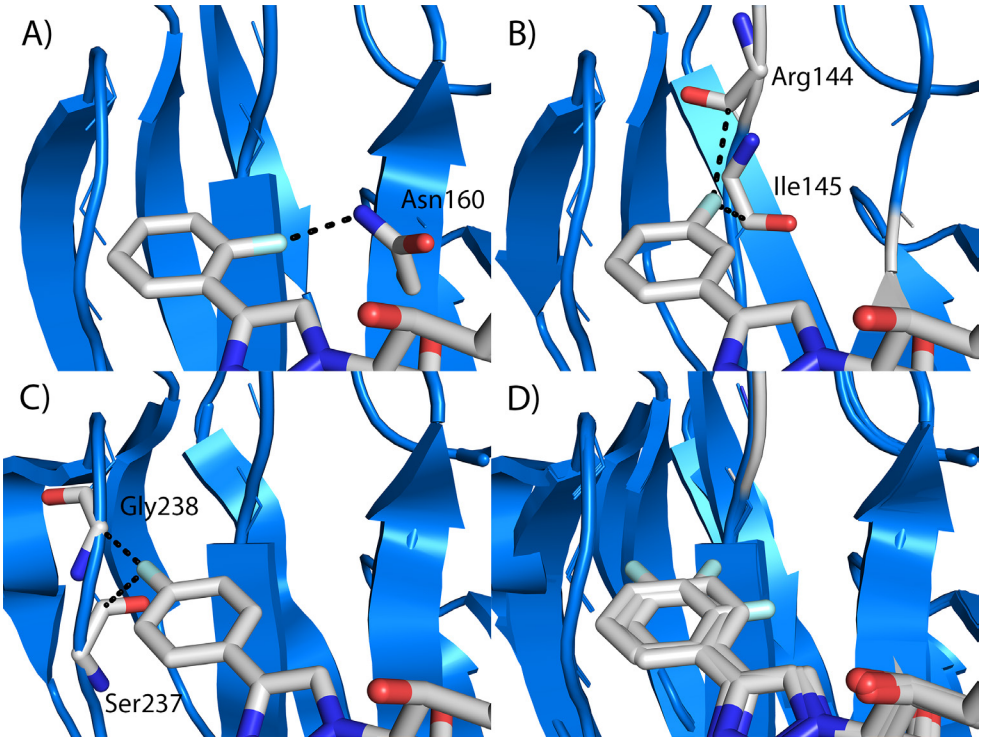


Figure 19. Close-up view of the fluorinated phenyl in the crystal structures of galectin-3C in complex with phenyltriazoles A) **17b**, B) **17c** and C) **17d**. D) Superimposed view of the three crystal structures. Observed fluorine interactions are indicated with black dashed lines.

The crystal structure of phenyltriazole **17g** (with fluorines at both meta and para positions) in complex with galectin-3C was also obtained (Figure 20). The phenyl ring of **17g** had a near complete overlap with the phenyl ring of 3-fluorophenyltriazole **17c**, which suggest that the orientation of the fluorine in meta position is more important than the fluorine in the para position. The orthogonal multipolar fluorine-amide interactions with Arg144, Ile145 and Ser237 and the multipolar C-F···H-C_α interaction with Gly238 observed in **17c-d** were also present in **17g**. The combination of these fluorine interactions explains the observed increase in galectin-3 affinity upon additional fluorination.

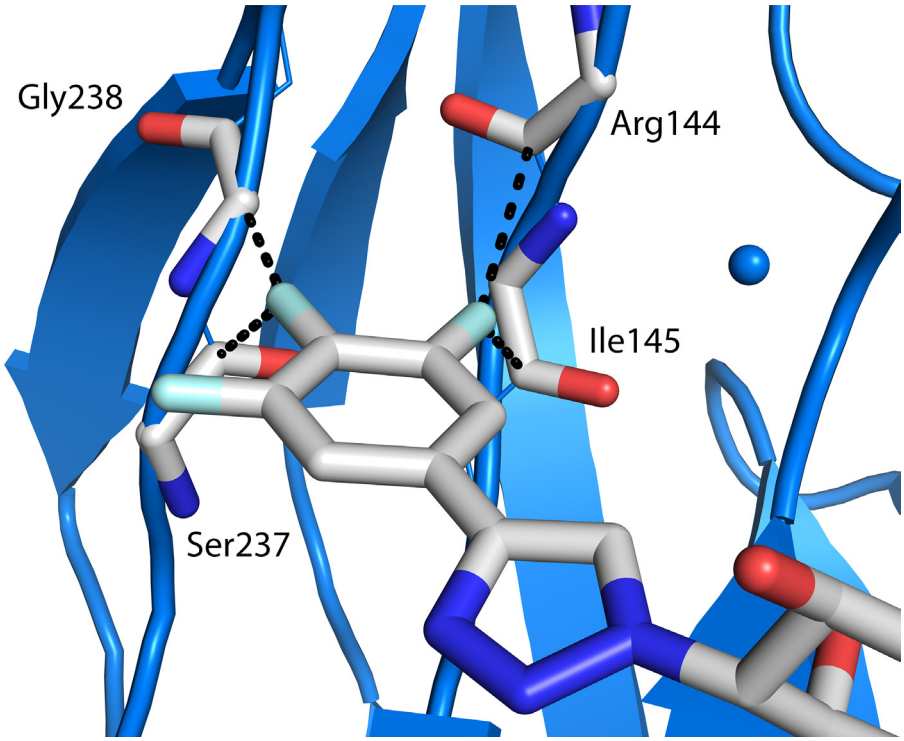


Figure 20. Close-up view of the trifluorophenyl ring in the crystal structure of trifluorophenyltriazole **17g** in complex with galectin-3C.

To further examine the observed fluorine interactions in phenyltriazoles **17b-d**, the more water-soluble galactosylthioglucosides **22a-c** were analyzed with ITC (Table 7, Figure 21). The ITC measurements revealed K_d values 4-fold higher than those obtained from FP for all ligands, as well as an enthalpy driven binding to galectin-3C. These mono-fluorinated phenyltriazoles were chosen for ITC analysis because their solvation was believed to be very similar. Although the negative entropic contributions were not identical, the main difference is the smaller enthalpic contribution when the fluorine is in the ortho position (**22a**) compared to meta- (**22b**) or para (**22c**) positions. This suggests fewer favorable binding interactions for **22a**, which further corroborates the lack of a favorable fluorine-amide interaction between the ortho fluorine (**22a**) and the side-chain amide of Asn160.

Table 7. Evaluation of galactosylthioglucosides **22a-c** with FP towards galectin-3 and direct titration ITC towards galectin-3C. Errors in parentheses showing a 68% confidence interval.

	K_d FP (μM)	K_d ITC (μM)	ΔG (kJ/mol)	ΔH (kJ/mol)	-TΔS (kJ/mol)
22a	1.9 ± 0.1	7.2 (6.0, 8.6)	-29.6 (-30.0, -29.2)	-45.5 (-47.5, -43.7)	15.9 (14.5, 17.5)
22b	0.46 ± 0.02	2.0 (1.7, 2.3)	-32.9 (-33.3, -32.5)	-50.4 (-51.6, -49.3)	17.5 (16.8, 18.3)
22c	0.60 ± 0.02	2.5 (2.3, 2.8)	-32.3 (-32.6, -32.1)	-49.1 (-49.9, -48.3)	16.8 (16.2, 17.3)

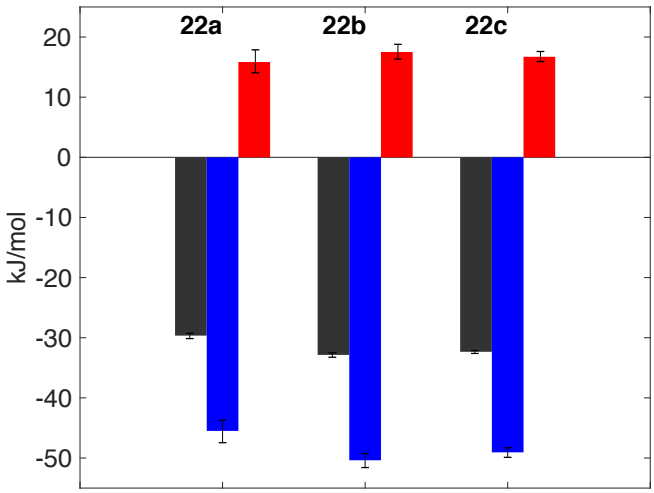


Figure 21. Thermodynamic parameters, ΔG (black bars), ΔH (blue bars) and -TΔS (red bars), of **22a-c** as determined in direct titration ITC experiments.

These fluorine interactions are the key to the 4-fold increase in galectin-3 affinity going from 3-fluorophenyltriazoles to 3,4,5-trifluorophenyltriazoles (*c.f.* **17c** and **17g** in Table 1 or **34a** and **34b** in Table 5). This affinity enhancement was further demonstrated when analyzing the crystal structures of asymmetric thiodigalactosides **33a-b** (a phenyltriazole at C3 and a coumaryl group at C3') in complex with galectin-3C (Figure 22). The 3-fluorophenyl analog (**33a**) bound galectin-3C with split occupancy, meaning the phenyltriazole and the coumaryl group forms approximately equally favorable interactions near Arg144 and Arg186. However, the 3,4,5-trifluorophenyl analog (**33b**) bound exclusively with the phenyltriazole in subsites A-B, which is likely due to the additional fluorine-amide interactions. In case of **33a**, Arg144 adopts a different conformation depending on the phenyltriazole or the coumaryl group binding in subsites A-B. However, in both cases Arg144 is stacked onto the π-system forming a cation-π interaction.

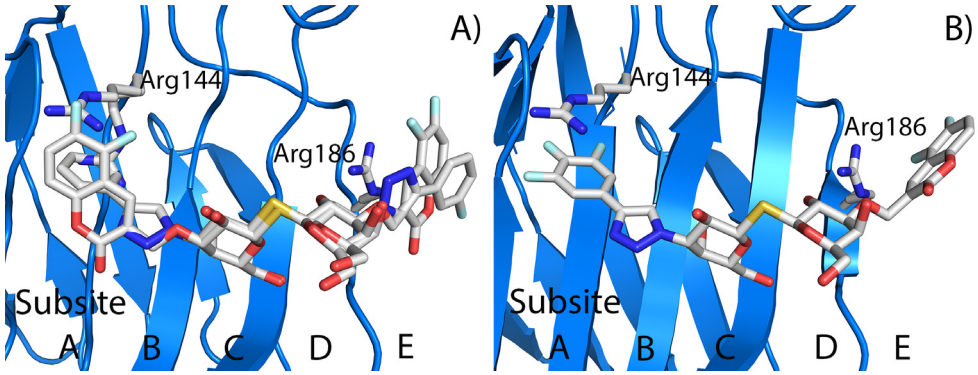
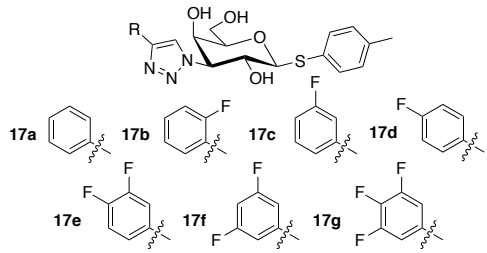


Figure 22. A) Crystal structure of thiodigalactoside **33a** in complex with galectin-3C revealing split occupancy of the ligand. B) Crystal structure of thiodigalactoside **33b** in complex with galectin-3C.

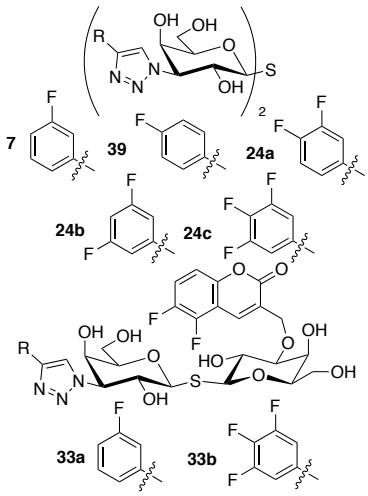
As mentioned earlier, thiodigalactosides with aromatic groups at C3 and C3' form cation-π interactions with Arg144 and Arg186. A cation-π interaction is stronger if the π-system is electron-rich, meaning it will weaken if electron-withdrawing fluorines are added to the π-system. Since the addition of fluorines to the π-system increase galectin-3 binding we decided to evaluate thiogalactosides **17a-g** towards galectin-3 R144S mutant (Arg144 is replaced with a serine), using FP (Table 8). The serine mutant was chosen to provide a minimal side-chain without introducing a non-polar surface. Overall, the effect of the R144S mutation was quite small suggesting that the stacking of Arg144 onto the phenyl ring does not contribute significantly to the binding. However, the surface underneath the phenyl ring remains hydrophobic in the R144S mutant and any interactions between the phenyl ring and this surface remains essentially unchanged. Interestingly, fluorination at the meta position (**17c**) seems more affected than the para position (**17d**). This is highlighted by the fact that **17c** binds galectin-3 R144S 7.3-fold weaker than galectin-3, while the drop in affinity for **17d** is only 1.6-fold. Replacing the arginine with a smaller serine may allow for increased rotation around the phenyl-triazole bond. This has a large effect on the meta fluorine-amide interactions where the fluorine is rotated out of position, while the para fluorine-amide interaction is not affected since rotation does not change the position of the fluorine. Based on these findings it seems that the cation-π interaction between the phenyl ring and Arg144 is rather weak and that Arg144 more importantly locks the phenyl in place, and its staggered conformation open the hydrophobic pocket that accommodates the fluorophenyl ring.

Table 8. K_d (μM) values and SEM of thiogalactosides **17a-g** towards galectin-3 and galectin-3 R144S.

	Galectin-3	Galectin-3 R144S	Galectin-3 R144S/ Galectin-3	Structures
17a	88 ± 3	250 ± 30	2.8	
17b	92 ± 5	250 ± 50	2.7	
17c	22 ± 0.7	160 ± 14	7.3	
17d	31 ± 1.3	51 ± 9	1.6	
17e	8.8 ± 0.3	35 ± 2	4.0	
17f	15 ± 0.3	74 ± 4	4.9	
17g	5.2 ± 0.3	38 ± 3	7.3	

Next, we decided to evaluate thiodigalactosides **6**, **7**, **24a-c**, **33a-b** and **39** towards galectin-3 R144S and galectin-3 R186S (Arg186 is replaced with a serine) to study the role of Arg186 and confirm the role of Arg144 using FP (Table 9). The galectin-3 R144S affinities show similar trends as was observed above with the meta-fluorinated phenyl **7** exhibiting a 5.2-fold decrease in affinity (*c.f.* galectin-3), while the para fluorinated phenyl **39** only exhibiting a 1.5-fold decrease. The galectin-3 R144S affinities of asymmetric thiodigalactosides **33a-b** can in part be explained by the increased rotation of the phenyl, but this should not by itself result in this large a difference. Another part of the explanation for the large decrease in galectin-3 R144S affinity of **33a** is found in the earlier mentioned two binding modes (Figure 22) of **33a-b** to galectin-3. Thiodigalactoside **33a** binds with split occupancy and when the coumaryl group is located in subsites A-B, Arg144 adopts a different conformation[26] forming a cation-π interaction. So with the coumaryl group of **33a** partially binding to Arg144, it is reasonable that its galectin-3 R144S affinity would be further decreased compared to **33b**. All thiodigalactosides bound galectin-3 R186S significantly weaker than galectin-3 indicating an important cation-π interaction between Arg186 and the ligand.

Table 9. K_d (nM) values and SEM for thiodigalactosides **6**, **7**, **24a-c**, **33a-b** and **39**.

	Galectin-3	Galectin-3 R144S	Galectin-3 R186S	Galectin-3 R144S/ Galectin-3	Structures
6	57 ± 4	300 ± 20	29000 ± 5200	5.3	
7	2.3 ± 0.2	12 ± 0.8	1000[27]	5.2	
39	4.0 ± 0.6	6.1 ± 1.0	1400 ± 70	1.5	
24a	1.1 ± 0.2	0.7 ± 0.2	830 ± 170	0.6	
24b	1.6 ± 0.3	2.9 ± 0.8	1100 ± 100	1.8	
24c	2.3 ± 0.4	1.9 ± 0.2	980 ± 150	0.8	
33a	12 ± 1	170 ± 10	20000 ± 3700	14.2	
33b	7.5 ± 0.7	19 ± 2	1900 ± 150	2.5	

In summary, three orthogonal multipolar fluorine-amide interactions with backbone amides (Arg144, Ile145 and Ser237) and a multipolar C-F···H-C_α interaction with Gly238 have been identified for meta- or para fluorinated phenyltriazoles (Figure 20), which significantly improve the affinity of those ligands. A fluorine-amide interaction with the side-chain amide of Asn160 was observed in the crystal structure, but could not be corroborated with thermodynamic- and affinity data. The cation-π interactions between aromatic groups at C3 and C3' of thiodigalactosides and Arg144 and Arg186 were investigated with two galectin-3 mutants (R144S and R186S). These thiodigalactosides were found to have a strong cation-π interaction to Arg186 and a weak cation-π interaction to Arg144.

5.2 Analysis of a deep pocket in subsite A of galectin-3 (Paper IV)

The synthesis of para substituted phenyltriazoles **18a-i** facilitated analysis of the deep pocket in subsite A of galectin-3 (Figure 10). The affinity of these ligands was discussed in section 4.3.1, which revealed a low tolerance for bulky substituents (**18f** and **18i**). Crystallization of phenyltriazoles **18a-e** and **18g-h** with galectin-3C was attempted, but due to a combination of low solubility and low affinity only crystal structures of **18a-d** (*p*-hydroxy, *p*-fluoride, *p*-azide and *p*-amine) and **18g** (*p*-methylamine) were obtained (Figure 23).

In the obtained crystal structures the ligands were positioned similarly, while Arg144 were found in two different conformations, either directly above the phenyl ring or above the para substituent. In some cases (fluoride **18b** and amine **18d**), split occupancy of the Arg144 was observed. The single occupancy of Arg144 in the galectin-3:**18g** complex is likely due to steric hindrance by the methyl group that points upward preventing Arg144 to be positioned above the NHMe group.

Considering a pKa value of 5.7[64] for pentafluorophenol (presumably similar to **18a**) it is possible that under the pH conditions used in FP (7.2) and X-ray crystallography (7.4) phenol **18a** is deprotonated when binding galectin-3. The deprotonation of **18a** results in more electron density around the oxygen and possibly a stronger interaction with Arg144. However, in the galectin-3:**18a** complex, we observe a displacement of the water between Asp239 and Arg144 that is present in the other complexes. The displacement of the water combined with the possibly stronger arginine-π interaction may explain the galectin-3 affinity of **18a**.

The Arg144 in the galectin-3:**18c** complex is positioned above the electron rich azide for a stronger cation-π interaction. The nitrogen connecting the azide to the phenyl ring is sp^2-hybridized placing the azide in the same plane as the phenyl ring allowing conjugation between the two π-systems. The phenyl-conjugated azide has two possible conformations, and it is pointing out in solution since pointing the other direction would result in a steric clash with Ile145.

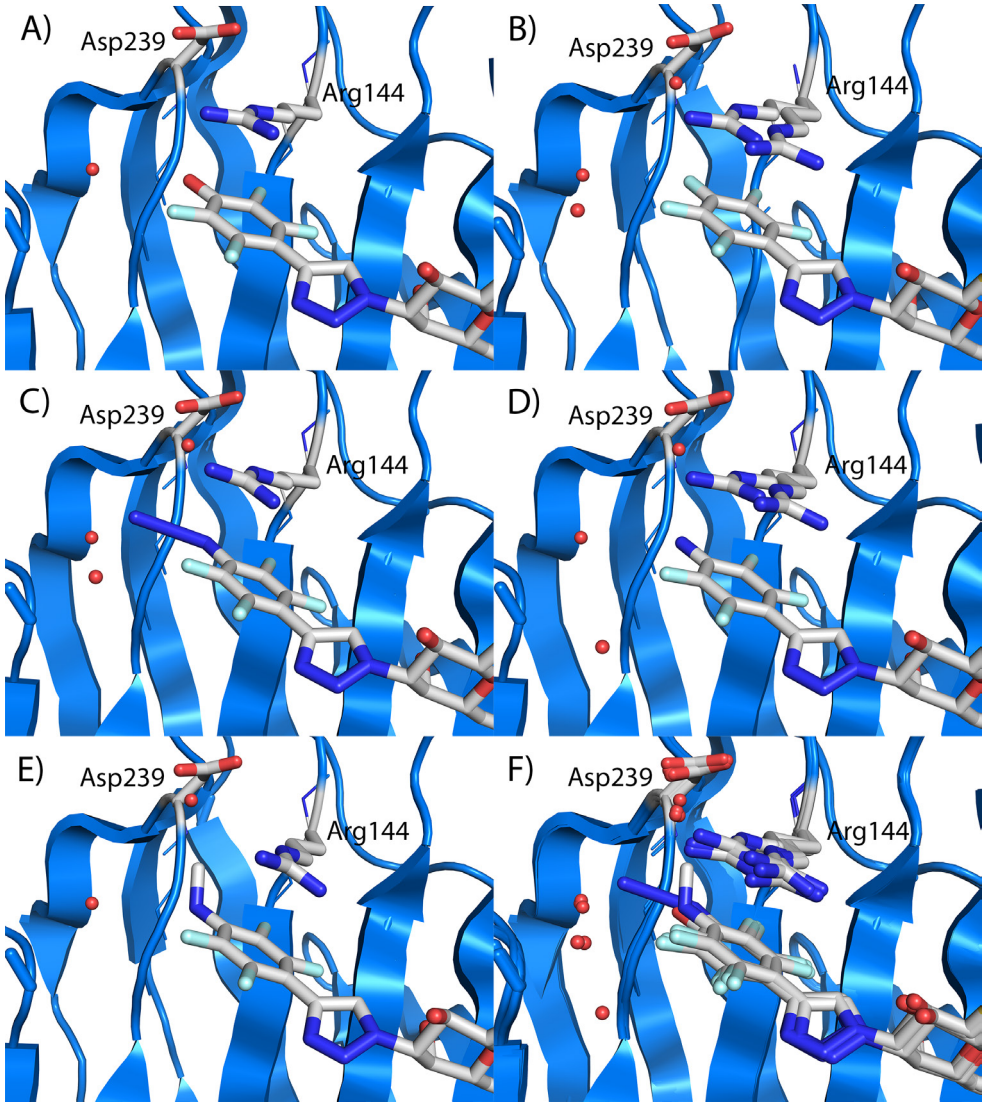


Figure 23. Close-up view of subsite A in the crystal structures of galectin-3C in complex with phenyltriazoles A) **18a**, B) **18b**, C) **18c**, D) **18d** and E) **18g**. F) Superimposed view of all crystal structures.

To summarize, the deep pocket in subsite A of galectin-3 does not seem to accommodate bulkier para substituents than –OMe or –NHMe, while electron rich para substituents *i.e.* azide seem to strengthen the cation-π interaction with Arg144.

5.3 Monosaccharide derivatives with low nM galectin-3 affinity (Paper VI)

The discovery of single-digit nM galectin-3 ligands is promising in the pursuit of potential drug candidates and thiodigalactoside **7** have successfully passed phase Ib/IIa clinical trial as a treatment for idiopathic lung fibrosis[65]. However, in order to increase membrane permeability we decided to make ligands with lower PSA, which was achieved by replacing the galactose unit in subsite D with a less polar aglycon. The starting point was one of our most potent (5.2 μM) monosaccharides, thiogalactoside **17g**, that has a 3,4,5-trifluorophenyltriazole at C3 and an equatorial 4-methylphenylthio at C1. We thus synthesized several ligands to optimize the thio-arene at C1 and discovered[66] a 16-fold increase in galectin-3 affinity when the thio-arene was placed axial instead of equatorial. Optimization of the substituents on the arene resulted in α-thiogalactoside **40** (Figure 24) with 37 nM[66] affinity towards galectin-3.

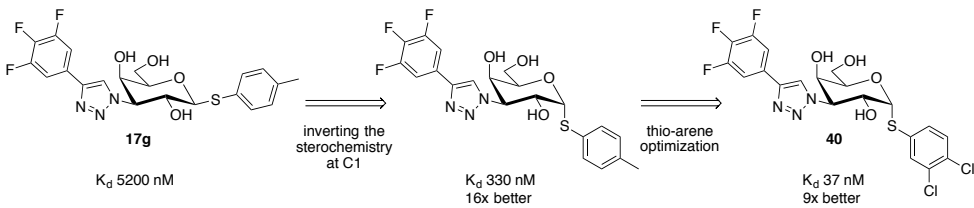


Figure 24. Schematic overview of the development of α-thiogalactoside **40**.

To explain the high affinity of α-thiogalactoside **40**, its complex with galectin-3C was analyzed with X-ray crystallography (Figure 25). The binding interactions in subsites A-B were similar to the previously discussed 3,4,5-trifluorophenyltriazole **17g** (Figure 20), while the axial placement of the thio-arene facilitated a sulfur-π interaction with Trp181. In subsite D, the σ-hole of the 3-chloro substituent was found to interact with a lone-pair of the carbonyl oxygen of Gly182 to form a halogen bond. These interactions would explain the remarkable nM affinity of α-thiogalactoside **40**.

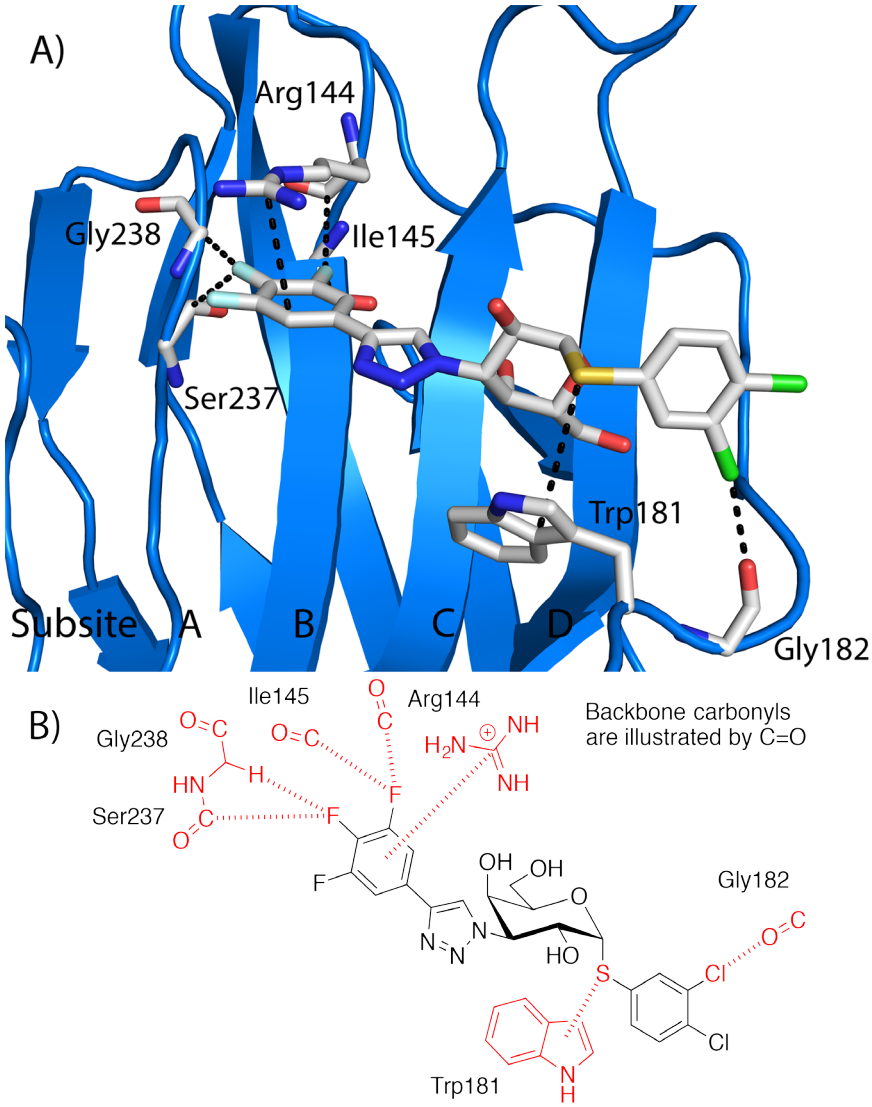
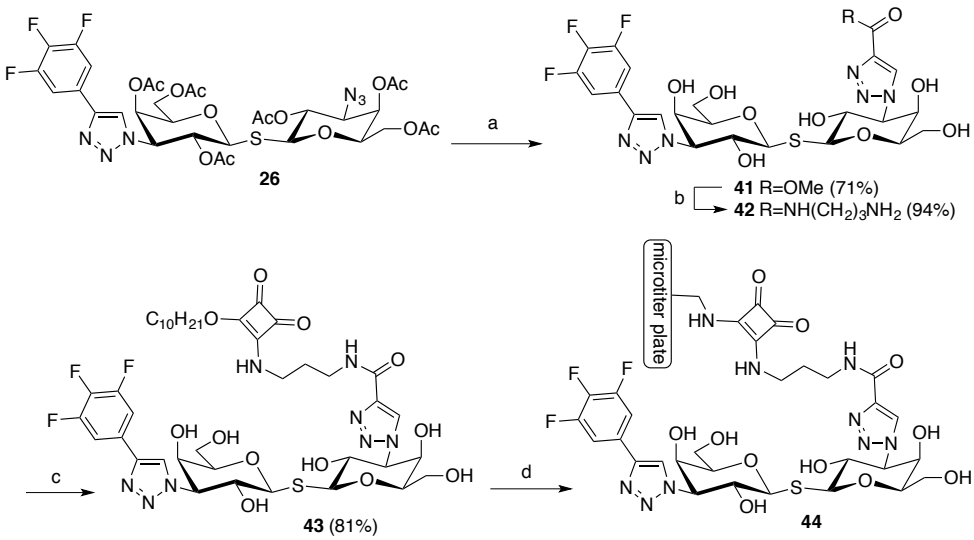


Figure 25. A) Crystal structure of α-thiogalactoside **40** in complex with galectin-3C. B) Overview of key interactions between α-thiogalactoside **40** and galectin-3.

6. A galectin-3 and 3C assay based on a high-affinity ligand immobilized in microwells (Paper VII)

Up-regulated expression of galectin-3 has been observed in patients with heart failure, and galectin-3 has thus been used as a biomarker for diagnosing heart failure.[67] Determination of galectin-3 levels in blood is currently performed with an enzyme-linked immunoassay[68] with a galectin-3 antibody that binds to the N-terminal of galectin-3, meaning it will only bind galectin-3, not galectin-3C. Since both galectin-3 and galectin-3C are present in blood we decided to develop a small-molecule based assay where a ligand with low nM affinity towards galectin-3 and galectin-3C is bound to a microtiter plate. This was envisioned by replacing the fluorescein group in probe **34b** with a squaric ester linker[69] that could be connected to an amino-functional CovaLink-NH microtiter plate (Scheme 10). Starting from thiodigalactoside **26**, amine **42** was synthesized via a 1,3-dipolar cycloaddition and deacetylation to give **41**, followed by addition of 1,3-diaminopropane. Using triethylamine as the base, selective monoaddition (**43**) to the squaric dodecyl ester was achieved. Switching to $NaHCO_3$ as the base, catalyzed the coupling of **43** to the microtiter plate resulting in ligand bound microtiter plate **44**. The remaining amino groups in the microtiter plate were capped using acetic anhydride in water.

Before evaluating galectin-binding to the microwells (**44**), the selectivity of the squaric ester **43** was determined towards different galectins using FP (Table 10). Squaric ester **43** bound galectin-3 and galectin-3C with 19 nM and 20 nM affinities, respectively, and **43** showed 40-10 000 fold selectivity compared to other tested galectins. The microplate **44** will, in due time, be evaluated towards all galectins but at this point only galectin-3 and galectin-3C have been tested. However, based on our knowledge of galectin-ligand binding, we believe the microplate (**44**) will exhibit similar galectin selectivity as the squaric ester (**43**). Both galectin-3 and galectin-3C were detected to bind strongly to **44** (Figure 26), thus confirming a successful immobilization of squaric ester **43** onto the microtiter plate.



Scheme 10. Synthesis of ligand bound microtiter plate **44**. Reagents and conditions: a) i. methyl propiolate, CuI, DIPEA, DMF, 50 °C; ii. NaOMe, MeOH, rt; b) 1,3-diaminopropane, MeOH, rt; c) 3,4-didecyloxy-cyclobut-3-ene-1,2-dione, Et_3N, DMF, rt; d) i. Amino-functional microtiter plate, $NaHCO_3$, DMSO, rt; ii. Ac_2O:H_2O (5:95), rt.

Table 10. K_d (μM) values and SEM for squaric ester **43** towards galectins-1, -2, -3, -3C, -4C, -4N, -8C, -8N, -9C and -9N.

Galectins									
1	2	3	3C	4C	4N	8C	8N	9C	9N
1.1 ± 0.2	0.94 ± 0.3	0.019 ± 0.0003	0.020 ± 0.0003	3.7 ± 0.6	2.9 ± 0.3	11 ± 2	190 ± 0.5	6.2 ± 0.6	15 ± 0.9

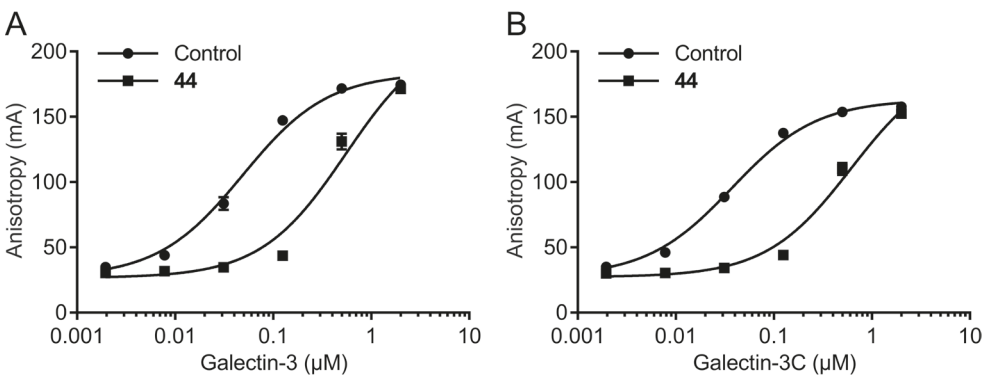


Figure 26. Fluorescence polarization titration of (A) galectin-3 and (B) galectin-3C with the fluorescent probe **34b** using assay **44** or control microplates (i.e. no ligand).

7. Concluding remarks and future perspectives

We set out to discover selective galectin ligands with high affinity towards galectin-1 and galectin-3, respectively, and decipher key binding interactions (Figure 27). In the course of discovering potent galectin ligands, a more robust synthetic route to 3-azido-3-deoxy-β-D-galactopyranosides, using stable aryl sulfonates was developed, which allows future scale-up.

An optimization of galectin binding in subsites A-B was performed through the synthesis of different aryltriazols at C3 of thiogalactosides, which revealed 2-thiazole as the optimal aryl for galectin-1 and 3,4,5-trifluorophenyl for galectin-3. Based on these findings, thiodigalactoside derivatives with single-digit nM affinity towards galectin-1 and galectin-3, respectively, were synthesized. New fluorescent probes and a protocol for competitive ITC were developed for better evaluation of high affinity galectin-1 and galectin-3 ligands.

Through the identification of three orthogonal multipolar fluorine-amide interactions with backbone amides and one multipolar C-F···H-C_α interaction, the potency and selectivity of galectin-3 ligands were improved. Utilizing these fluorine-amide interactions and the discovery of an important halogen bond in subsite D, monosaccharides with low nM galectin-3 affinity were developed.

In the future, our discovered galectin-1 and galectin-3 ligands could, directly or after optimization, be used to investigate biological functions and treat diseases. Our insights into galectin-ligand binding would facilitate the design and synthesis of new ligands towards other galectins. Galectin-3 has worked well as a model protein for investigation of biomolecular recognition and we envision that more fundamental thermodynamic-, structural- and computational analyses will be performed on galectin-3.

The galectin-3 affinity-enhancing 3,4,5-trifluorophenyl moiety was exploited in the development of small-molecule based assay for detection of galectin-3 and galectin-3C, by the successful immobilization of the small-molecule ligand to the microtiter plate. Further experiments are required to reach the end-goal of this assay, which is to be able to accurately determine galectin-3 and galectin-3C levels in blood serum.

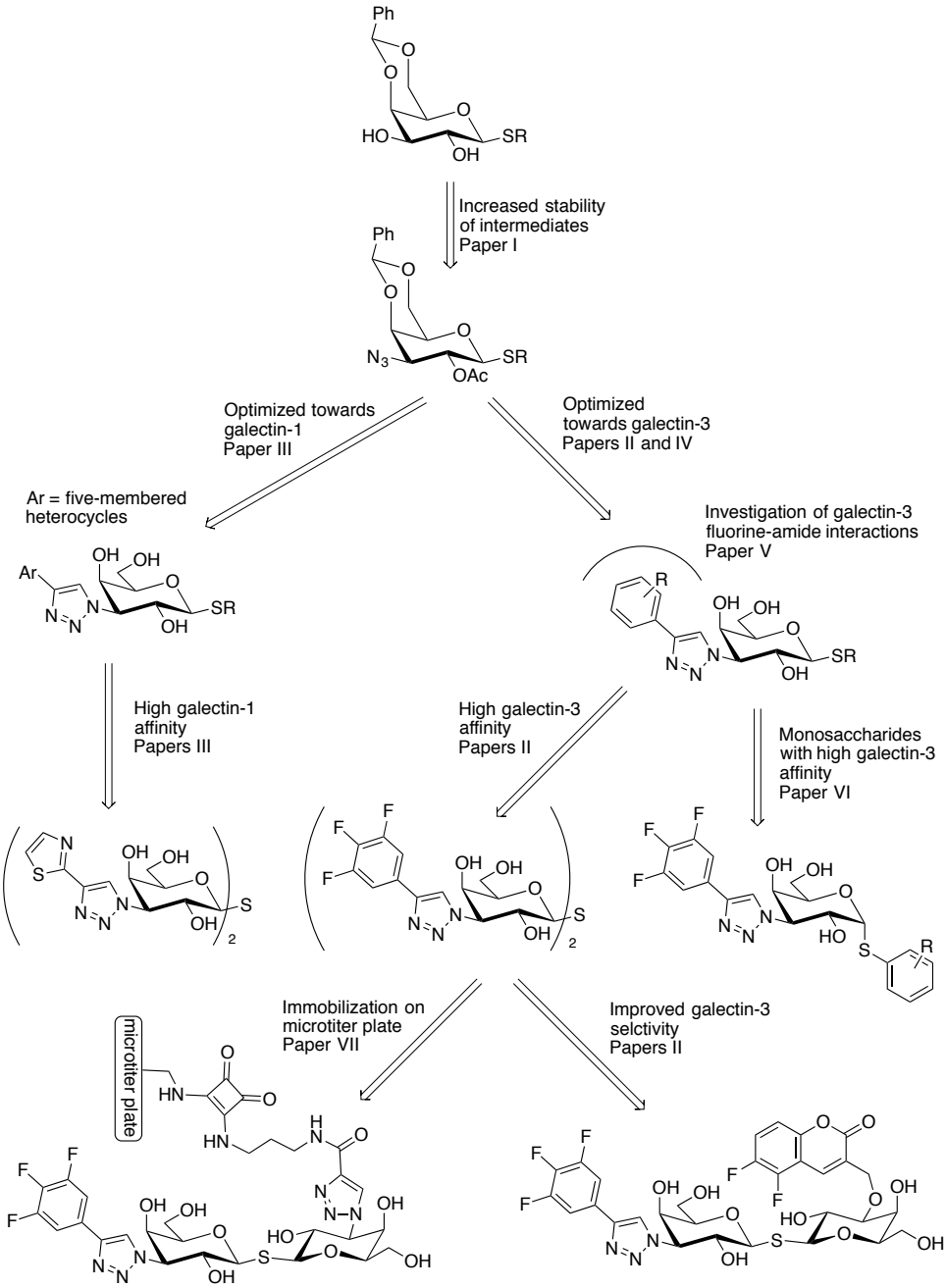


Figure 27. Summary of all projects.

8. Acknowledgements

The past 5 years of my life as a PhD student at Lund University have been an exciting and rewarding experience and I am very grateful for the opportunity I was given. Completing this thesis would not have been possible without the financial support of Lund University, the Swedish Research Council, Galecto Biotech, Knut and Alice Wallenberg foundation, the Royal Physiographic Society, the Royal Swedish Society of Science. There are a lot of people that have helped me along the way that I would like to thank.

Ulf Nilsson, min handledare, för att du välkomnade mig i din forskargrupp och all kunskap du bidragit med. Det har varit en fantastisk tid med massor av spännande idéer och projekt. Ingenting är omöjligt med din positiva inställning och optimism.

Ulf Ellervik, min andra Ulf, en outtömlig inspirationskälla.

Fredrik Zetterberg, för all kunskap och diskussioner om utveckling av läkemedel.

Hakon Leffler och *Barbro Kahl-Knutsson*, för allt som berör galektiner och utvärdering av galektinhämmare, utan er hade galektinprojektet inte fungerat.

Current and former members of the Nilsson Group, for the great times in the lab, on conferences and excursions. I will miss you all.

All colleagues in the DecRec research network, for all the help with experiments and great scientific discussions.

All the people who helped co-author the papers presented in this thesis.

Maria, Katarina och Bodil, utan er hade forskningen på CAS inte fungerat.

Sofia, Gösta och Maggan, för all pedagogisk assistans och service med MS och LC.

Karl-Erik och Anders, för all support med NMR och IT.

Alla seniorer på CAS, för all kunskap ni bidrar med.

All colleagues at CAS, for making the great working environment at CAS.

Mamma och pappa, Sofia och Johan, för att ni alltid stöttat mig i allt jag tagit mig vid och alltid funnits där för mig. Jag är enormt tacksam för allt ni gjort för mig. Jag är den jag är tack vare er.

References

1. Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683–720.
2. Barondes, S. H.; Castronovo, V.; Cooper, D. N. W.; Cummings, R. D.; Drickamer, K.; Felzi, T.; Gitt, M. A.; Hirabayashi, J.; Hughes, C.; Kasai, K.; Leffler, H.; Liu, F.-T.; Lotan, R.; Mercurio, A. M.; Monsigny, M.; Pillai, S.; Poirer, F.; Raz, A.; Rigby, P. W. J.; Rini, J. M.; Wang, J. L. *Cell* **1994**, *76*, 597–598.
3. Leffler, H.; Carlsson, S.; Hedlund, M.; Qian, Y.; Poirier, F. *Glycoconj. J.* **2002**, *19*, 433–440.
4. Nieminen, J.; Kuno, A.; Hirabayashi, J.; Sato, S. *J. Biol. Chem.* **2007**, *282*, 1374–1383.
5. Seetharaman, J.; Kanigsberg, A.; Slaaby, R.; Leffler, H.; Barondes, S. H.; Rini, J. M. *J. Biol. Chem.* **1998**, *273*, 13047–13052.
6. Rabinovich, G. a; Rubinstein, N.; Fainboim, L. *J. Leukoc. Biol.* **2002**, *71*, 741–752.
7. Liu, F.-T.; Rabinovich, G. A. *Nat. Rev. Cancer* **2005**, *5*, 29–41.
8. Boscher, C.; Dennis, J. W.; Nabi, I. R. *Curr. Opin. Cell Biol.* **2011**, *23*, 383–392.
9. Thijssen, V. L.; Heusschen, R.; Caers, J.; Griffioen, A. W. *Biochim. Biophys. Acta Rev. Cancer* **2015**, *1855*, 235–247.
10. Astorgues-Xerri, L.; Riveiro, M. E.; Tijeras-Raballand, A.; Serova, M.; Neuzillet, C.; Albert, S.; Raymond, E.; Faivre, S. *Cancer Treat. Rev.* **2014**, *40*, 307–319.
11. Liu, F.-T.; Rabinovich, G. A. *Ann. N.Y. Acad. Sci.* **2010**, *1183*, 158–182.
12. MacKinnon, A. C.; Gibbons, M. A.; Farnworth, S. L.; Leffler, H.; Nilsson, U. J.; Delaine, T.; Simpson, A. J.; Forbes, S. J.; Hirani, N.; Gauldie, J.; Sethi, T. *Am. J. Respir. Crit. Care Med.* **2012**, *185*, 537–546.
13. Vrasidas, I.; André, S.; Valentini, P.; Böck, C.; Lensch, M.; Kaltner, H.; Liskamp, R. M. J.; Gabius, H.-J.; Pieters, R. J. *Org. Biomol. Chem.* **2003**, *1*, 803–810.
14. Salameh, B. A.; Leffler, H.; Nilsson, U. J. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3344–3346.
15. Giguère, D.; Bonin, M. A.; Cloutier, P.; Patnam, R.; St-Pierre, C.; Sato, S.; Roy, R. *Bioorg. Med. Chem.* **2008**, *16*, 7811–7823.

16. Tejler, J.; Salameh, B.; Leffler, H.; Nilsson, U. J. *Org. Biomol. Chem.* **2009**, *7*, 3982–3990.

17. Sörme, P.; Qian, Y.; Nyholm, P. G.; Leffler, H.; Nilsson, U. J. *ChemBioChem* **2002**, *3*, 183–189.

18. Sörme, P.; Arnoux, P.; Kahl-Knutsson, B.; Leffler, H.; Rini, J. M.; Nilsson, U. J. *J. Am. Chem. Soc.* **2005**, *127*, 1737–1743.

19. Glinsky, V. V.; Kiriakova, G.; Glinskii, O. V.; Mossine, V. V.; Mawhinney, T. P.; Turk, J. R.; Glinskii, A. B.; Huxley, V. H.; Price, J. E.; Glinsky, G. V. *Neoplasia* **2009**, *11*, 901–909.

20. Rauthu, S. R.; Shiao, T. C.; André, S.; Miller, M. C.; Madej, É.; Mayo, K. H.; Gabius, H. J.; Roy, R. *ChemBioChem* **2015**, *16*, 126–139.

21. Dion, J.; Deshayes, F.; Storozhylova, N.; Advedissian, T.; Lambert, A.; Viguier, M.; Tellier, C.; Dussouy, C.; Poirier, F.; Grandjean, C. *ChemBioChem* **2017**, *18*, 782–789.

22. Cumpstey, I.; Sundin, A.; Leffler, H.; Nilsson, U. J. *Angew. Chem. Int. Ed.* **2005**, *44*, 5110–5112.

23. Delaine, T.; Cumpstey, I.; Ingrassia, L.; Mercier, M. Le; Okechukwu, P.; Leffler, H.; Kiss, R.; Nilsson, U. J. *J. Med. Chem.* **2008**, 8109–8114.

24. Cumpstey, I.; Salomonsson, E.; Sundin, A.; Leffler, H.; Nilsson, U. J. *Chemistry* **2008**, *14*, 4233–4245.

25. Salameh, B. A.; Cumpstey, I.; Sundin, A.; Leffler, H.; Nilsson, U. J. *Bioorg. Med. Chem.* **2010**, *18*, 5367–5378.

26. Rajput, V. K.; MacKinnon, A.; Mandal, S.; Collins, P.; Blanchard, H.; Leffler, H.; Sethi, T.; Schambye, H.; Mukhopadhyay, B.; Nilsson, U. J. *J. Med. Chem.* **2016**, *59*, 8141–8147.

27. Delaine, T.; Collins, P.; MacKinnon, A.; Sharma, G.; Stegmayr, J.; Rajput, V. K.; Mandal, S.; Cumpstey, I.; Larumbe, A.; Salameh, B. A.; Kahl-Knutsson, B.; van Hattum, H.; van Scherpenzeel, M.; Pieters, R. J.; Sethi, T.; Schambye, H.; Oredsson, S.; Leffler, H.; Blanchard, H.; Nilsson, U. J. *ChemBioChem* **2016**, *17*, 1759–1770.

28. Öberg, C. T.; Leffler, H.; Nilsson, U. J. *J. Med. Chem.* **2008**, *51*, 2297–2301.

29. Öberg, C. T.; Blanchard, H.; Leffler, H.; Nilsson, U. J. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3691–3694.

30. Cumpstey, I.; Carlsson, S.; Leffler, H.; Nilsson, U. J. *Org. Biomol. Chem.* **2005**, *3*, 1922–1932.

31. Sörme, P.; Kahl-Knutsson, B.; Wellmar, U.; Magnusson, B. G.; Leffler, H.; Nilsson, U. J. *Methods Enzymol.* **2003**, *363*, 157–169.

32. Nilsson, U.; Leffler, H.; Mukhopadhyay, B.; Rajput, V. Novel Galactoside Inhibitors of Galectins. WO2013110704 A1, Aug 1, **2013**.

33. Bissantz, C.; Kuhn, B.; Stahl, M. *J. Med. Chem.* **2010**, *53*, 5061–5084.

34. Muller, K.; Faeh, C.; Diederich, F. *Science* **2007**, *317*, 1881–1886.

35. Zürcher, M.; Diederich, F. *J. Org. Chem.* **2008**, *73*, 4345–4361.

36. Hof, F.; Scofield, D. M.; Schweizer, W. B.; Diederich, F. *Angew. Chem. Int. Ed.* **2004**, *43*, 5056–5059.

37. Cavallo, G.; Metrangolo, P.; Milani, R.; Pilati, T.; Priimagi, A.; Resnati, G.; Terraneo, G. *Chem. Rev.* **2016**, *116*, 2478–2601.

38. Salonen, L. M.; Ellermann, M.; Diederich, F. *Angew. Chem. Int. Ed.* **2011**, *50*, 4808–4842.

39. Beno, B. R.; Yeung, K.-S.; Bartberger, M. D.; Pennington, L. D.; Meanwell, N. A. *J. Med. Chem.* **2015**, *58*, 4383–4438.

40. Patrick, G. L. *An Introduction to Medicinal Chemistry*; Fourth Edi.; Oxford University Press, **2009**.

41. Renaud, J.-P.; Chung, C.; Danielson, U. H.; Egner, U.; Hennig, M.; Hubbard, R. E.; Nar, H. *Nat. Rev. Drug Discov.* **2016**, *15*, 679–698.

42. Edwards, M. P.; Price, D. A. In *Annual Reports in Medicinal Chemistry*; Elsevier Inc., **2010**; Vol. 45, pp. 380–391.

43. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3–26.

44. Pajouhesh, H.; Lenz, G. R. *NeuroRX* **2005**, *2*, 541–553.

45. Leeson, P. D.; Springthorpe, B. *Nat. Rev. Drug Discov.* **2007**, *6*, 881–890.

46. Cheng Li; Tongtong Liu; Xiaoming Cui; Uss, A. S.; Cheng, K.-C. *J. Biomol. Screen.* **2007**, *12*, 1084–1091.

47. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.

48. Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057–3064.

49. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599.

50. Worrell, B. T.; Malik, J. A.; Fokin, V. V. *Science* **2013**, *340*, 457–460.

51. Öberg, C. T.; Noresson, A.-L.; Delaine, T.; Larumbe, A.; Tejler, J.; von Wachenfeldt, H.; Nilsson, U. J. *Carbohydr. Res.* **2009**, *344*, 1282–1284.

52. Muramatsu, W. *J. Org. Chem.* **2012**, *77*, 8083–8091.

53. Vatèle, J.-M.; Hanessian, S. *Tetrahedron* **1996**, *52*, 10557–10568.

54. Lowary, T. L.; Hindsgaul, O. *Carbohydr. Res.* **1994**, *251*, 33–67.

55. Peterson, K.; Kumar, R.; Stenström, O.; Verma, P.; Verma, P. R.; Håkansson, M.; Kahl-Knutsson, B.; Zetterberg, F.; Leffler, H.; Akke, M.; Logan, D. T.; Nilsson, U. J. *J. Med. Chem.* **2018**, acs.jmedchem.7b01626.

56. van Scherpenzeel, M.; Moret, E. E.; Ballell, L.; Liskamp, R. M. J.; Nilsson, U. J.; Leffler, H.; Pieters, R. J. *ChemBioChem* **2009**, *10*, 1724–1733.

57. Sörme, P.; Kahl-Knutsson, B.; Huflejt, M.; Nilsson, U. J.; Leffler, H. *Anal. Biochem.* **2004**, *334*, 36–47.

58. Checovich, W. J.; Bolger, R. E.; Burke, T. *Nature* **1995**, *375*, 254–256.

59. Pierce, M. M.; Raman, C. S.; Nall, B. T. *Methods* **1999**, *19*, 213–221.

60. Leavitt, S.; Freire, E. *Curr. Opin. Struct. Biol.* **2001**, *11*, 560–566.

61. Krantz, M. J.; Holtzman, N. A.; Stowell, C. P.; Lee, Y. C.; Weiner, J. W.; Liu, H. H. *Biochemistry* **1976**, *15*, 3963–3968.

62. Hsieh, T.-J.; Lin, H.-Y.; Tu, Z.; Lin, T.-C.; Wu, S.-C.; Tseng, Y.-Y.; Liu, F.-T.; Hsu, S.-T. D.; Lin, C.-H. *Sci. Rep.* **2016**, *6*, 29457.

63. Bum-Erdene, K.; Gagarinov, I. A.; Collins, P. M.; Winger, M.; Pearson, A. G.; Wilson, J. C.; Leffler, H.; Nilsson, U. J.; Grice, I. D.; Blanchard, H. *ChemBioChem* **2013**, *14*, 1331–1342.

64. Kraut, D. A.; Sigala, P. A.; Pybus, B.; Liu, C. W.; Ringe, D.; Petsko, G. A.; Herschlag, D. *PLoS Biol.* **2006**, *4*, e99.

65. Galecto Biotech. https://galecto.com (accessed 20 Jan, 2018).

66. Zetterberg, F. R.; Peterson, K.; Johnsson, R. E.; Brimert, T.; Håkansson, M.; Logan, D. T.; Leffler, H.; Nilsson, U. J. *ChemMedChem* **2018**, *13*, 133–137.

67. Yin, Q.-S.; Shi, B.; Dong, L.; Bi, L. *J. Geriatr. Cardiol.* **2014**, *11*, 79–82.

68. Milting, H.; Ellinghaus, P.; Seewald, M.; Cakar, H.; Bohms, B.; Kassner, A.; Körfer, R.; Klein, M.; Krahn, T.; Kruska, L.; El Banayosy, A.; Kramer, F. *J. Heart Lung Transplant.* **2008**, *27*, 589–596.

69. Bergh, A.; Magnusson, B.-G.; Ohlsson, J.; Wellmar, U.; Nilsson, U. J. *Glycoconjugate J.* **2001**, *18*, 615–621.